DOCUMENT RESUME

ED 154 830 IR 005 982

AUTHOR	Richards, William D., Jr.
TITLE	A Coherent Systems Methodology for the Analysis of Human Communication Systems. Report No. 25.
INSTITUTION	Stanford Univ., Calif.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Mar 76
GRANT	DCR73-07973, A02
NOTE	278p.
EDRS PRICE	MF-$0.83 HC-$15.39 Plus Postage.
DESCRIPTORS	*Algorithms; *Communication (Thought Transfer); Doctoral Theses; Group Dynamics; Intercommunication; *Mathematical Models; Matrices; *Models; Networks; *Research Methodology; *Systems Approach
IDENTIFIERS	GASSICS

ABSTRACT

The first part of this four-part dissertation reviews conceptual and operational approaches to human communication systems. A new conceptual model for studying communication, General Algorithm for Studying Structuring in Complex Systems (GASSICS), is developed in the second part, and an operational method for applying GASSICS is presented in part three. Part four includes an example of the GASSICS method of analysis, and an examination of the potential of this method in aiding the advancement of theory and the guidance of research in the general area of human communication systems. (Author/CMV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED154830

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A COHERENT SYSTEMS METHODOLOGY FOR THE ANALYSIS OF HUMAN COMMUNICATION SYSTEMS

William D. Richards, Jr.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Edwin B. Parker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

March 1976

Report No. 25

National Science Foundation Grant DCR73-07973, A02

PROGRAM IN INFORMATION TECHNOLOGY AND TELECOMMUNICATIONS
CENTER FOR INTERDISCIPLINARY RESEARCH
Stanford University Stanford, California 94305

(A dissertation submitted to the Department of Communication and the Committee on Graduate Studies of Stanford University in Partial fulfillment of the requirements for the degree of Doctor of Philosophy).

IR005982

ACKNOWLEDGMENTS

This dissertation was completed with the guidance of Edwin B. Parker, who prevented me from straying too far from my chosen direction, and who provided me with encouragement when it seemed that the effort demanded by theoretical work would not be justified by the results. I am grateful also to Everett M. Rogers and Gerhard J. Hanneman, who provided much useful feedback, and to Richard V. Farace, whose support has extended from the earliest beginnings to the present. I also thank the students who took classes from me while I was working out these ideas. They put up with a lot of unfocussed discussion and helped me to see more clearly what I wanted to say. Special thanks to my friends who remained friends through the many months of ups and downs: I hope I will be as patient with them as they were with me. Finally, I gratefully acknowledge the support of National Institute of Health Biomedical Communication and Information Science Training Grant (# 5 TO 1 LM00135-04) and National Science Foundation Economics of Computer Communication Networks Grant (# DCR73-07973, A02), which supported much of the writing of this dissertation.

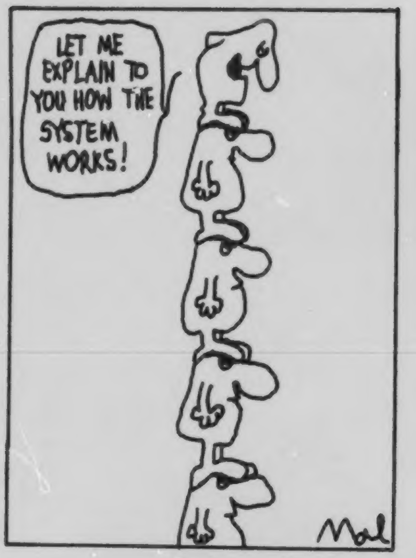

LET ME EXPLAIN TO YOU HOW THE SYSTEM WORKS!
Mal

TABLE OF CONTENTS

List of Figures x

List ot Tables xii

PART I: INTRODUCTION 1

CHAPTER ONE: THE MECHANISTIC AND ORGANISMIC MODELS OF REALITY 4

- THE MECHANISTIC MODEL AND THE CLASSICAL ANALYTIC MODEL 7
- CLASSICAL MECHANISMS 8
- THE MECHANISM FAILS: DIALECTIC 10
- DIALECTIC: SYNTHESIS 13
- THE SYSTEMS MODEL 15

CHAPTER TWO: PREVIOUS METHODS OF NETWORK ANALYSIS 20

- INTRODUCTION 20
- THE SOCIOGRAM 20
- MATRIX METHODS 22
  1. Distance vs. Linkage Models 23
  2. The Definition of Group 24
  3. The Method of Identifying Group 25
- THE METHODS 26
  1. Matrix Manipulation 26
  2. Matrix Multiplication 26
  3. Factor Analysis 28
  4. Multidimensional Scaling Methods 28
  5. Blockmodelling Methods 29
  6. Miscellaneous Methods 31
- Problems 31
  1. Matrix Manipulation 33
  2. Matrix Multiplication 33
  3. Factor Analysis 33
  4. Multidimensional Scaling Methods 34
  5. Blockmodelling Methods 34
  6. Miscellaneous Methods 35
- Relative Advantages 35
- ON MATRIX REPRESENTATIONS 35
- EXTANT DATA 37
- CONCLUSION 38

PART TWO: CONCEPTUALIZATION 40

CHAPTER THREE: THE FORM OF SYSTEMS 43

INTRODUCTION 43
1. Interaction 47
2. Levels 48
3. Interactions Between Systems: The Form/Logic of Information 52

CONCLUSION 58

CHAPTER FOUR: THE PROCESSES OF DESCRIPTION AND OBSERVATION: ACCESS TO FORM 59

1. The Form of Descriptions 59
2. The Process of Observation 64
3. Constraints on the Available Data 67
4. Dealing with Constraints -- Tools 71

CHAPTER FIVE: STRUCTURE: THE APPEARANCE OF FORM 73

REVIEW 73

PREVIEW 77

STRUCTURE 78
Structure and Randomness 78
Emergent Properties, Levels, and Structures 79
Structure and Form 80
Properties, Descriptions, and Constraints 82
Sources of Constraint: Structure vs. Control Systems . . 84
Summary 85

STRATEGIES FOR RESEARCH 86

A GENERAL ALGORITHM FOR STUDYING STRUCTURE IN COMPLEX SYSTEMS 88
The Loop 89

PART THREE: OPERATIONALIZATIONS 92

CHAPTER SIX: STEP ONE OF GASSICS: DATA COLLECTION 94

INTRODUCTION 94
Assumptions 94
Implications of the Assumption of Discrete Levels 97

RELATIONSHIPS 97
General Characteristics of Relationships 98
1. Strength 98
2. Symmetry 98
3. Transitivity 98
Matter-energy vs. Information Relationships 98

OPERATIONALIZATION IN GENERAL 99
Symmetry/Reciprocity 100

Strengths 102
Transitivity 102

HUMAN COMMUNICATION NETWORKS 104
General Concepts 104
    1. Content or Mode 104
    2. Strength 105
    3. Reciprocity/Symmetry 106
Actual Instrumentation Considerations 108
    1. Scaling 109
    2. Other Instrumentation Considerations 111
An Alternate Operationalization of Strength 113
Diaries: An Alternative to Recall Methods 115
Higher Levels in Communication Networks 116

CHAPTER SEVEN: STEP TWO OF THE GASSICS PROCEDURE: THE MEASUREMENT OF STRUCTURING 118

Preliminary Discussion 119

OPERATIONALIZATION 120
"Triangles" and Constraints 123
The Expected Number of Triangles: The Random Model . . . 128
Variance in the Distribution of Links 129
Correcting T_e for the Effects of ℓ_i 130
Maximum Limits to T and S^2 135
Relative Structure Measures 137

SYSTEM-LEVEL IMPLICATIONS OF STRUCTURING 139

SUMMARY 144

CHAPTER EIGHT: THE IDENTIFICATION OF DIFFERENTIATED SUBUNITS 146

INTRODUCTION 146
Terminology 146
Goals of G-Network Analysis 147
    1. Linkage vs. Distance Representation 147
    2. The Classification of Nodes 148

AN ALGORITHM FOR PERFORMING G-NETWORK ANALYSIS 155
The Classification of Non-participants 156
The Algorithm Which Identifies Groups 156
    1. The Vector Averaging Process 158
    2. Drawing the Tentative Boundaries 166
    3. Using the Criteria for an Exact Solution 174

CHAPTER NINE: THE SYSTEMS APPROACH, GASSICS, AND GNA 184

THREE LEVELS OF SPECIFICITY 184
    1. The General Systems Approach 184
    2. The GASSICS Procedure 187
    3. The Actual Operations 188

THE DEVELOPMENT OF THE ANALYSIS PROCEDURES 190
- 1. G-Network Analysis: Four Breakthroughs 190
  - a. NETWOW: vector averaging 190
  - b. "Continuous Matrices" 191
  - c. The Group Identification Algorithm 196
  - d. The Exact Criteria 199
- 2. Data Collection Methods: All Links are not Equal 200
  - a. Nominal Binary Approaches 200
  - b. Ordinal Strength Indicators 201
  - c. Ratio-level Indicators 201
- 3. Structural Indicators 202

PART IV: RESEARCH 204

CHAPTER TEN: USING G-NETOWRK ANALYSIS IN RESEARCH 204

INTRODUCTION 204

REVIEW OF SELECTED NETWORKS 205
- 1. The Chase Studies 205
- 2. The OCD Studies 206
- 3. The Treasure Island Studies 206
- 4. The Naval Studies 207

Discussion 208

AN EXAMPLE OF NETWORK ANALYSIS: THE MELBOURNE STUDY . . 209
- 1. Background 209
- 2. Structural Analysis 211
- 3. The Network Analysis 211
  - a. Basic Findings 213
  - b. Gross Comparisons with other Networks . . . 213
  - c. Non-participants 213
  - d. Liaisons and Others 218
  - e. Group Members 219
  - f. Bridges and Non-bridges 219
  - g. Individuals and their Links 219
  - h. Groups in the Network 222
    - 1) How many and how large 222
    - 2) Links in groups 225
    - 3) Group size, group density, and $\overline{D}$ 225
    - 4) Outside-group connections 229
  - i. The System 230

RELATING NETWORK VARIABLES TO OTHER VARIABLES 233
- 1. Structural Stability over Time 233
- 2. Antecedents and Consequences of Network Roles . 239
- 3. The Impact of Group Structure on Attitudinal Configurations 240

SUMMARY 242

CHAPTER ELEVEN: FURTHER 244

INTRODUCTION 244

THE STATE OF THE ART 246

WHERE WE ARE/WHERE WE WANT TO GO 247
Theory 247
Statistical and Methodological Considerations 250
    1. Sampling 250
    2. Measurement 251
    3. Descriptive Statistics 253
Analytic Considerations 254
    1. Improvements in Existing Programs 254
    2. Development of Other Programs 255
    3. Moving Towards an "NPSS" (Network Package for the Social Sciences) 255

DISCUSSION 255

CONCLUSION 258

Bibliography 259

APPENDIX 266

LIST OF FIGURES

CHAPTER TWO

Figure 1: A Sociogram 21
2: Imaginary Data Illustrating the Matrix Manipulation Procedure 27
3: Imaginary Data Illustrating Blockmodels, Lean Fit, and Zeroblocks 30

CHAPTER FIVE

Figure 1: GASSICS: A General Algorithm for Studying Structuring in Complex Systems 91

CHAPTER SEVEN

Figure 1: Constraint and the Organization of Links . . . 122
2: Interconnection and the Number of Triangles . . 124
3: Maximizing the Number of Triangles With Group Structure 126
4: Minimizing the Number of Triangles With Group Structure 127
5a: Plot of Expected variance Against Number of Links 131
5b: Plot of Expected Variance Against System Density 131
6: Generalized Plot of Expected Variance Against System Density 132
7: The Relation Between Observed Variance and Number of Triangles 133
8: Maximizing the Number of Triangles by Reducing Effective N 136
9: Generalized Plot of Both Maximum Number of Triangles and Expected Number of Triangles Against System Density 138
10: The Relation Between Differentiation into Groups and Number of Triangles 140

CHAPTER EIGHT

Figure 1a: The Network Roles Illustrated 150
1b: Classification Tree of Network Roles 151
2: The Identification of Non-participants 157
3: The "Balls and Bands" Model 159
4a: The Vector Averaging Process: A Hypothetical Network 162

Figure 4b: The Vector Averaging Process: Computing Means 163
5: The Iterative Process of Vector Averaging . . . 164
6: A Hypothetical Network and its Continuum . . . 167
7: The Window-scanning Process and the Construction of the Density Plot 170
8: Cutting Points and the Boundary-drawing Process 172
9: A Distance Matrix 176
10: The Distance Matrix and Critical Nodes 179
11: Flowchart for the G-Network Analysis Algorithm. 183

CHAPTER NINE

Figure 1: Four Representations of a Hypotethical Network: (a) as a Sociogram; (b) as a Discrete Binary Adjacency Matrix; (c) as a "continuous" Binary Adjacency Matrix; and (d) as a Continuum . . . 193
2: Different Ways of Constructing the Density Plot 198

CHAPTER TEN

Figure 1: The Relation Between Group Size and Group Density in the Melbourne Network 226
2: The Relation Between Group Size and $\overline{D}$ for all Groups in the Melbourne Study 227
3: The Relation Between Group Size and $\overline{D}$ for all Groups in the Melbourne Study 228
4: The Relation Between Distance to Other Group Members and Number of Outside-group Links 231
5: A "Sociogram" for the Groups in the Melbourne Network, Drawn from the Matrix Shown in Figure 10 234

LIST OF TABLES

CHAPTER 7

Table 1a: The ℓ_i's for Three Sample Networks Shown in Figure 7-7 134
1b: Structural Values for Three Sample Networks Shown in Figure 7-7 134
2a: The ℓ_i's for Three Sample Networks Shown in Figure 7-10 141
2b: Structural Values for the Sample Networks Shown in Figure 7-10 141

CHAPTER 10

Table 1: Structural Analysis of the Melbourne Network . . 212
2: The Basic Findings of the Melbourne Study 214
3: Breakdown of All Networks by Roles 215
4: The Melbourne Network: The Links of Participants 220
5: Mean Number of Links by Roles Across Networks . . 221
6: Cross-role Connections -- the Melbourne Network: Breakdwon of All Links by Role 223
7: Cross-network Comparison of Group Structure . . . 224
8: Structural Analysis of the Melbourne Network Using Groups Instead of Nodes 232
9: Comparison of Participant and Isolation Proportions of 14 Networks 236
10: Comparison of Roles Across Time within Network: The Naval Networks 238
11: Stability Across Content and Time: The Naval Networks 238

PART I

INTRODUCTION

One of the most pervasive and basic characteristics of the world we live in is change. Current times are characterized by rates of change unprecedented in recorded history. Everything is changing, faster and faster all the time. Electronic devices are getting smaller and smaller, the population is growing larger and larger, and problems of all kinds are becoming more and more complex.

Many of these changes are reflective of a basic shift in the kinds of processes engaged in by people living in many parts of the world -- a shift in communication style. This shift is fed by the continued development and advancement of new communication technologies. Thus, easy access to relatively inexpensive telephone equipment increases the amount of communication between distant areas, in much the same way railroads, and later airplanes, increased the amount of communication by making rapid mail service possible.

Recent research (Parker, 1975) suggests that we are becoming an information society, where the primary commodity that is processed is information, rather than the industrial matter and energy that characterized the pre-information society.

Not only is there more communication, but also is there faster communication. Where it used to take weeks to get a letter from San Francisco to New York, it is now possible for most of the world to watch men walking on the surface of the moon, with a delay dictated only by the speed of light.

As educational levels rise and political barriers drop, more and more people gain the ability to interact in the context of the emerging world society. Where in the past, local and national societies were forced to be independent of one another by a lack of communication facilities, they are now tied together into what is fast becoming a single integrated network of interdependent units, where the boundaries are becoming more and more mere political or economic considerations, instead of natural geographical or racial barriers.

What is the result of this increased interdependence? Gerard (1968, p. 53) says that "the more the members of the class interact ... rather than coexist, the more does the superordinate group become a true individual rather than a collection of ordinate individuals." Instead of a world of separate, independent nations, we are moving in the direction of a single world society, composed of smaller interacting national or regional groups.

This shift is not evident only on a global scale. The tendency in the American economic system has been to move away from mechanical processes and toward information processes. We are becoming more and more reliant on information, as the total volume of information exchanged in a given unit of time is growing rapidly as communication and information technologies advance. One implication of this growth is a greater interdependence among the segments of the political-economic system. Cases where the effects of an apparently local change reverberate through the entire system are being seen with increasing regularity. It is obvious that the addition of more communication links is changing the fundamental nature of the system.

The nature of this change is becoming a legitimate topic for study from a number of viewpoints, as more and more problems are being recognized as system problems which can only be understood in the context of the larger system. This new viewpoint has been effective in a large number of traditional fields -- economics, environmental studies, transportation, and education, to name a few. In spite of the vast differences in their topic areas, there seems to be a growing concensus that researchers and theoreticians feel a need to understand the general properties of complex systems.

This need is being translated into research. The "systems approach" is a phrase heard with increasing regularity at professional meetings and seen more and more in journals and, recently, textbooks. An unmistakable sign of the arrival of the systems approach is its occurrence in popular fiction. There is even a novel about self-reproducing man-made systems, with a title that suggests one of the potential dangers of uncontrolled complex systems -- Mechasm (Sladek, 1968).

There have been many papers with the phrases "A Systems Analysis of" or "A Systems Approach to" in their title. There is a large body of literature made of uses or applications of systems thinking. There is also a body of writing about the systems approach in general, theoretical terms. There is much less, however, in the way of systems methodologies. This dissertation presents an attempt to bridge the gap from systems theory to systems research. The bridge takes the form of a research approach — a methodology, a guiding paradigm that structures and leads the research endeavor.

The dissertation is divided into four parts. The first reviews prior conceptual and operational approaches to the problems. A new conceptualization is developed in the second part, and a new operational method is presented in the third. There are two chapters in the fourth part. In the first we present an example of the new method of analysis. In the second the new method is examined for its potential to aid in the advancement of theory and the guidance of research in the general area.

CHAPTER ONE

THE MECHANISTIC AND ORGANISMIC MODELS OF REALITY

> "Thus the classical picture 'works' where it is applicable. And because of its success it has a terrible seductive power over the human mind." (Rapoport and Horvath, 1968, p. 72)

> " ... is not the understanding of any complexity to be gained by analyzing it further and further into its constituent parts?" (Rapoport and Horvath, 1968, p. 72)

> "All the king's horses and all the king's men couldn't put Humpty together again." (Mother Goose)

Until only recently man's approaches to the study of man have been based upon two major conceptual models -- the mechanism (the mechanical device which obeys the laws of physics) and the organism (the living, growing, evolving plant or animal). Buckley notes that "sociological theory has been living for some time off the intellectual capital of previous centuries" (1967, p. 1). He goes on to suggest that this dependence on the traditional conceptual structures is the basis for the "sizeable chorus of critics and skeptics" of the great bulk of empirical research conducted in the last decades. The difficulty, he says, "lies in the fact that current dominant theory is built on mechanical and organic (more exactly, organismic) systems models laid down during previous centuries" and which are quite inappropriate in dealing with the kinds of systems we are interested in (1967, p. 1).

Deutsch (1968) argues that the use of these models in attempting to understand social situations was consistent with advances in the physical and natural sciences, where these models were very successful. According to Monge:

> The breakthrough by the natural sciences in developing viable conceptualizations inspired social scientists to adopt models based on the natural science conceptualizations of the world, all of which was done in hopes of achieving similar success in explaining human behavior; thus, physical and biological science success bred social science imitation. Unfortunately, the hoped-for success has never been realized; social and communication science are still without a viable model of human behavior (1973, p. 7).

Of course, there have been many models proposed for the communication process since the seventeenth and nineteenth centuries, when the mechanistic and organismic models were first applied to the study of man. These include the Lasswell model of "Who says What in Which Channel to Whom with What Effect" (1948); and the Shannon and Weaver model of 1949, as adapted by Berlo in his SMCR model (1960), which focused on the source, message, channel, and receiver (in parallel to Lasswell's model). McCroskey (1968) added a feedback loop, and Dance (1967) drew the model as a helix instead of a circle, to suggest the passage of time.

Monge suggests that a review of other models, such as the Westley-MacLean (1957) conceptual model, the Becker (1968) mosaic model, and the Barnlund (1970) transactional model, among others, would be ...

> ... fruitless, for despite their proliferation, virtually all contemporary theories of communication lack the sophistication necessary to be classified as even mechanistic or organismic, much less as systems models. Rather, they tend to be pictorial or verbal descriptions of a poorly conceptualized vaguely defined phenomenon (1973, p. 10).

If we examine the literature of the social sciences, we would see that the vast bulk of the research that has been done has centered around the individual. In the relatively much smaller volume of literature that is concerned with social systems (which may be as small as a simple two-person dyad or as large as an entire civilization), the mechanical and organismic models of reality are basic to most of the research that has been done.

This is not to say that researchers consciously used mechanical or organismic analogies as, for example, Woelfel (1970) did in his "Theory of Linear Force Aggregation," where messages and the self-concept are related by Newton's Second Law which equates force with the product of mass and acceleration. Rather, social scientists incorporated into their thinking at a very basic level certain fundamental assumptions concerning "the way things are" consistent with those upon which the mechanistic and organismic models were based. It is thus the use of these basic assumptions that identifies the later research as fundamentally mechanistic or organismic. Both of these basic approaches are severely limiting in that:

(1) They restrict the kinds of situations or phenomena that can be identified and utilized in scientific explanation, because they fail to consider adequately the complexities of the processes of observation and description.

(2) They restrict the kinds of logical "moves" that can be made in going from raw data to final interpretive statement, because they fail to recognize the relation of data and descriptions to the reality being studied.

(3) They thus restrict the kinds of theory that can be put forth, both because of the expressive difficulties mentioned in (1) and because of the strategic logical difficulties mentioned in (2). Furthermore, the simplistic nature of these approaches practically guarantees a parallel simplicity in theory building, as the basic models contain no hints of the kinds of complexity that can be seen in hierarchical information processing systems, and thus do not suggest how these systems should be approached.

Let us review the two basic models and see how they lead to the difficulties mentioned above.

THE MECHANISTIC MODEL AND THE CLASSICAL ANALYTIC MODEL

Within the current dominant paradigm, scientific assertions take the form of "if so ... then so" statements, which are usually interpreted as assertions of causality. Indeed, "the prediction paradigm of science interprets all scientific assertions as assertions of valid causality relations" (Rapoport, 1968, p. xiii). Mathematical physics, a particularly successful and highly developed branch of natural science, uses a more powerful type of assertion than the causality statement -- the equation -- in its descriptions of dynamic relationships between entities across time. In the mathematical language of systems of equations, a single statement will embody an infinity of "if so ... then so" statements, as it is expressed in continuous, rather than discrete, terms. In addition, in the language of mathematics, it is possible to efficiently deal with situations where there are complex and even recursive networks of causal relationships, all interacting with one another over time.

> The connection between the language of equations and the 'vulgate' language of causality is established by holding constant all of the variables except one pair. This enables us to say: 'Other things being equal, the thinner the bottom of the kettle, the sooner the water will boil'; or 'Other things being equal, the greater the atmospheric pressure, the longer it will take to bring the water to a boil,' etc. Thus, 'common-sense' causal relations are included in the equation and are deduced by holding constant all the variables except those of interest.
>
> In assuming that the equation in which all the causal factors were combined was given, we have, of course, assumed that all of the relations were known at once. In actuality they are often determined one by one. These separate determinations are made possible by the method of controlled experiment. In order to bring out some causal relation free of disturbances by other factors, we deliberately try to hold constant all those factors suspected of having some influence. Thus the basic assumption underlying the empirical study of physical phenomena is that we can eliminate all disturbing phenomena and study the relation of interest alone. Next, by establishing several pairs of such relations, we can (we assume) combine them into a more general

> causality law, that is, an equation in which all the contributing factors appear as variables. This is called the analytic method. It has been phenomenally successful in the physical sciences (Rapoport, 1968, p. xiv).
>
> Analysis is an attempt to understand a complexity by examining its constituent parts. The parts being simpler, they are supposedly more amenable to understanding. The idea of analysis, then, is to understand the working of the parts .... The implied hope is that it is possible to 'build up' the understanding of a complexity by 'superimposing' the workings of the various parts (Rapoport and Horvath, 1968, p. 87).

The central ideas here, namely (1) that we can eliminate all disturbing phenomena and study the relation of interest alone, and (2) that we can combine descriptions of pairs of relations established by studying isolated aspects of the situation into more general causality laws, are unquestionably mechanistic in tone. Recall the status of scientific thought when the classical "mechanism" first led to successful prediction of future events with Newton's description of the solar system.

According to Deutsch (1968, p. 388), mechanistic analogies were quickly applied to descriptions of government by Hobbes, Locke, Montesquieu, and de la Mettrie. Tom Paine extended them to God as the "first mechanic," and Schiller spoke of the "watchspring of the universe." This extension of the idea of mechanism, from the experience of newly developed pumps and clockworks, says Deutsch (1968, p. 388), to a general description of reality, was encouraged by Newton's success with mechanical descriptions of gravitational astronomy.

CLASSICAL MECHANISMS

> The physical concepts of space, time, attraction, inertia, force, power -- which must be recognized as anthropomorphisms originally borrowed from everyday human experience -- were borrowed back in their new connotative attire and applied to man and society. Thus we find conceptions of moral or

> *social space* in which social events occur; *position* in social space, and a *system of social coordinates* defining man's position in it; social processes as results of the 'gravitation' or *attraction* and *inertia* of individuals and groups, the latter regarded as a *system in an equilibrium* of *centrifugal* and *centrifugal forces*. Social organization, power, and authority were resultants of the *'pressures'* of 'social *atoms*' and 'molecules': hence arose 'social statics' or a theory of social equilibrium analogous to statics in physical mechanics, and 'social dynamics' involving motion or change as a function of time and space expressible by various mathematical curves (Buckley, 1967, p. 8).

> The classical concept or model of mechanism implied the notion of a whole which was completely equal to the sum of its parts; which could be run in reverse; and which would behave in exactly identical fashion no matter how often those parts were disassembled and put together again, and irrespective of the sequence in which the disassembling or reassembling would take place. It implied consequently the notion that the parts were never significantly modified by each other, nor by their own past, and that each part once placed in its appropriate position, with its appropriate momentum, would stay exactly there and continue to fulfill its completely and uniquely determined function (Deutsch, 1968, p. 388).

It is easy to see how the classical analytic method is well suited to the study of mechanisms, and why it should be chosen to guide the study of a reality thought to behave as a mechanism. In fact, the mechanistic model was very successful in the physical sciences, and the analytic method was the one that allowed this success. As long as the systems being studied were fairly simple, and as long as the measurement techniques remained relatively insensitive to discrepancies between the reality and the model, the analytic technique worked well: it was possible to take the more complex situations apart into simpler situations that could be easily understood. The results of this piecemeal analysis could then be combined to give a more or less adequate description of the whole. Adequate, that is, until the reality being studied failed to fit the mechanistic model to such an extent that the discrepancies could not be ignored.

THE MECHANISM FAILS: DIALECTIC

As machines became more complex, so did the sets of interrelationships among their parts. The effects of the past history of machines became appreciable, as the greater complexity resulted in machines whose performance was dependent on precise intermeshing of many parts, which were subject to wear. Furthermore, attempts to apply this approach to the study of animals and, in particular, man were met with disappointment more often than success. In the face of repeated failures, the first assumption made by researchers was probably that their measurement or analysis techniques were imprecise. It was easier to improve these areas than to construct a new model of reality, and so these were the areas where improvements were made (witness the current state of statistical methods — complex regression techniques and sophisticated multidimensional scaling routines).

In some areas of investigation, the old methods in their extended versions seemed to work. In others, however, they continued to fail. The classical analytic method would not work when the situation under investigation involved closed causal loops, as, for example, did all systems with operating feedback loops. This was especially the case with living processes.

> Views on why this is so differed. According to one view, called vitalism, the extension is in principle impossible, because living processes are not governed by the same laws as nonliving processes. According to the opposite views, variously called mechanism, physicalism, or reductionism, the difficulty lies not in an irreducible difference between physical and biological laws but only in the tremendous complexity of living processes. The reductionists assumed implicitly that if we knew enough about how living beings were put together, we could write down the equations that govern their behavior; and if we were clever enough in mathematics, we could solve the equations and so determine the 'trajectories' of behavior (Rapoport, 1968, p. xv).

There were thus two responses to the failure of the mechanistic model. The mechanists said essentially that their model was still adequate.

More careful and extensive analysis (in the classical sense) of complex systems (such as animals and man) was all that was needed. The controlled experiment was still the preferred investigatory strategy.

The organists, on the other hand, said that in living systems the whole approach taken by the mechanists was doomed to failure at the start, because organisms are not mechanisms -- they are fundamentally different, and this difference demands a different approach for understanding. It was almost as if all of reality were divided into two mutually exclusive classes — mechanisms and organisms.

> According to the classical view, an 'organism' is unanalyzable, at least in part. It cannot be taken apart and put together again without damage. As Wordsworth put it, 'We murder to dissect.' The parts of a classical organism, insofar as they can be identified at all, not only retain the functions which they have been assigned but in fact cannot be put to any other functions (except within narrow limits of 'de-differentiation' which were often ignored), without destroying the organism. The classical organism's behavior is irreversible. It has a significant past and a history -- two things which the classical mechanism lacks — but it is only half historical because it was believed to follow its own peculiar 'organic law' which governs its birth, maturity, and death and cannot be analyzed in terms of clearly identifiable 'mechanical' causes (Deutsch, 1968, p. 389).

The classical organismic approach thus seems to deny all the assumptions of the mechanistic viewpoint. A purely organismic research strategy would do the opposite of the mechanistic approach. Where a mechanist would take apart a complex system to see how its parts worked in hopes of gaining an understanding of the whole, the organist would deny this as a viable strategy. The reason for a given system's behavior would be "the system" -- the origins of behavior are to be found in the wholeness of the system, and are not reducible to the constituent parts of the system. Otherwise, the organism would be reducible to a mechanism, and this state of affairs would be very hard to accept as a model of one organism in particular -- man.

For this reason mechanistic explanations of organisms would not be expected to "work." However, there were difficulties with mechanical explanations even in the physical sciences. Attempts to find mechanical models for quantum mechanics, for example, were not successful, and were "taken as evidence for the 'mysterious' character of subatomic processes" and for the existence of a "pervasive 'spiritual reality' that is not indifferent or alien to human values" (Nagel, 1961, p. 337). Thus, signs of organism were evident, even in the domain of mathematical physics, where the mechanical model had been most successful:

> ... the failure to explain electromagnetic phenomena in terms of mechanics, and the general decline in mechanics from its earlier position as the universal science of nature, have been construed as evidence for the 'bankruptcy' of classical physics, for the necessity of introducing 'organismic' categories of explanation in the study of all natural phenomena, and for a variety of sweeping doctrines concerning levels of being, emergence, and creative novelty (Nagel, 1961, p. 337).

Thus the attribution of mysterious "life forces" or "higher forms" or different logical structures to an area that is resistant to modelling with mechanistic approaches seems to be common, even in the physical sciences.

Although modern organismic biases are still with us, they seldom speak of "life forces" or "entelechies." Instead, they refer to "organic wholeness" or to "functional unity."

> It follows that understanding cannot be extended beyond the scope of physical science without introducing concepts which embody irreducible wholes in place of physically measurable variables. The concept of organism is indispensible in biology; the concept of the individual in psychology; the concepts of the institution and social class in sociology; the concept of a nation in contemporary political science; the concept of a culture in anthropology. Each of these wholes presents itself naturally, because we perceive it as such (Rapoport, 1968, p. xvii).

Of course our thinking is guided by our perceptions. Thus it seems logical to study these complex wholes as we perceive them -- that is, as individual units, which we "murder to dissect." This approach is valid, in that it does respect the oneness of complex systems. However, attempts to apply organismic thinking to gain an understanding -- not only of what a complex system does as a system, but also of how the system does what it does, given that it is what it is -- are not often fruitful.

> We recognize an organism, an individual, a nation; and we assume that in proper circumstances it acts as a whole. Still, if we confined our attention exclusively to the grossly observable patterns of these wholes, we would not make much progress toward understanding this behavior (Rapoport, 1968, p. xvii).

Just as the mechanistic approach ignored too much by assuming a system could be studied piecemeal, so did the organismic approach prevent understanding by denying the reducibility of complex phenomena to the interaction of constituent components.

DIALECTIC: SYNTHESIS

The mechanistic approach is opposed by the organismic in a dialectic where fundamental views of reality are contested. Where the organismic approach resulted from the opposition of the mechanistic method to situations not suited to mechanical analysis, the systems approach comes from the new dialectic.

De la Mettrie, writing in 1747, suggested a way out of the animate vs. inanimate (i.e., mechanism vs. organism) dilemma with his suggestion that ....

> .... Matter was in itself neither organic nor inorganic, neither living nor dead, neither sensible nor insensible. The difference between these states or properties of material things sprang, not from the intrinsic natures of their raw materials, but from the different ways in which these materials were organized (in Toulmin and Goodfield, 1962, p. 318).

Although De la Mettrie had no suggestions on how to study systems, he got very close to what has turned out to be a crucial concept in modern systems theory — organization. The way things are organized -- how one part is related to another in a context — has a profound influence on relationships the thing may enter. This is especially important when the "things" are complex systems and the relationships are ones of observation and description.

The dialectic is not yet fully developed. However, we can see how the prior steps are limiting, as we look from a broader perspective than the ones provided by either of the two basic models.

(1) Both restrict the kinds of situations or phenomena that can be identified and utilized in scientific explanation by failing to consider adequately the complexities of the processes of observation and description: the mechanistic approach looks at systems as simple collections or aggregates of parts, while the organismic approach "sees" only in terms of wholes. Neither recognizes the subjective nature of the perceptual process, which includes the arbitrary (often unconscious) imposition of distinctions which allow us to speak and think of "units." Once the arbitrary nature of these distinctions is recognized, the importance of understanding the relationship between a unit at one level (a part, perhaps) and a unit at another level (the whole, perhaps) becomes manifest. This process must become conscious if a priori conceptualizations are not to influence the processes of observation and description in such a way that we are effectively blinded by our subjective perceptions. Perhaps a more conscious approach, one which recognizes the subjective a priori nature of perceptual processes, will allow us to develop more useful, less restricting priors.

(2) Both restrict the kinds of logical "moves" that can be made in going from raw data to final interpretive statements, by failing to recognize the relation of data and description to the reality being studied. On one hand the mechanistic approach assumes that complex phenomena can be taken apart and studied piecemeal. The bits of knowledge are then added together to give an understanding of the whole. Since this approach ignores the effects due to mutual multiple interactions among the parts, it cannot anticipate the effects of these interactions, thus limiting the kinds of statements that can be made, given a set of raw data. This limitation stems from the

combination of two factors: (a) when moving up to a higher level of analysis, it is assumed that the information about the lower levels is simply added up to provide a description of the system at the higher level — it is not necessary to provide a different description for the different level; and (b) not only is it not necessary to provide a different description (at a different level of analysis) for the different level, but there is no way to generate such a description, given the original one.

On the other hand, the organismic approach denies the legitimacy of cross-level analysis as even a logical possibility. Thus, data must be used at the level from which they come. For example, the relation of the behavior of the individual in an organization to the behavior of the organization as a unit in its own right is not open to scientific investigation.

(3) The limitations discussed above combine to restrict the kinds of theory that can be put forth. The basic problem seems to stem from a fundamental ignorance of relationships between levels — levels of analysis, levels of abstraction, levels of perception, levels of description. When in the study of complex, multi-leveled systems these relationships are violated, confused, or ignored, it becomes difficult to move toward an understanding. When the relationships are clarified, research can progress much faster, as important questions are both more readily identified and resolved. This will translate into more substantive theory, together with a fundamentally deeper understanding.

THE SYSTEMS MODEL

> A whole which functions as a whole by virtue of the interdependence of its parts is called a system, and the method which aims at discovering how this is brought about in the widest variety of systems has been called general system theory. General system theory seeks to classify systems by the way their components are organized (interrelated) and to derive 'laws,' or typical patterns of behavior, for the different classes of systems singled out by the taxonomy (Rapoport, 1968, p. xvii).

> A system is a set of objects together with relationships between the objects and between their attributes (Hall and Fagan, 1968, p. 18).

The whole is more than the sum of the parts.

The systems approach, like the organismic approach, says that we have to look at the whole system. If we isolate the parts, we take them out of the context in which they function, thus destroying their relationship to the system or, in other words, changing the system. Unlike the organismic approach, however, the systems approach identifies the parts, together with their interrelationships, as the origins of properties seen in the whole system. The systems approach denies the existence of "life forces" -- or else it reduces them to effects of the interaction of the parts of the system.

One example of this process is seen in the reduction of "purpose" or "goal-oriented" behavior to cybernetic control systems, governed by negative feedback, where the discrepancy between the actual state of the system and the desired or ideal or goal state is fed back into the system in such a way as to cause it to move toward the goal state. With this arrangement, mindless machines can be made which appear to function as if they had conscious purposes.

The concepts of systems theory have enjoyed wide use by communication scientists for the past 10 or 15 years. Terms like "feedback," "boundary," and "process" appear often in the literature. At professional meetings we hear the phrase "systems approach" with increasing regularity.

But what does it mean to take a "systems approach" in communication research? To many, it means that we try to "get the whole picture" — look at all the relationships. This is sometimes translated into highly elaborate multivariate research strategies, i.e., "ask more questions." Any number of statistical methods may be used to analyze the results: metric multidimensional scaling techniques, multiple regression methods, or path analysis techniques, to name a few. But, some would argue, these are merely extensions of the mechanistic (i.e., classical analytic) model. They are not really different — they are not really systems approaches.

Buckley (1967, pp. 36-7) suggests that it is the "central focus on the principle of organization per se, regardless of what it is that is organized" that makes the systems approach. He goes on to say that we should find the systems approach attractive because it promises to develop:

> (1) A common vocabulary unifying the several 'behavioral' disciplines;
>
> (2) A technique for treating large, complex organization;
>
> (3) A synthetic approach where piecemeal analysis is not possible due to the intricate interrelationships of parts that cannot be treated out of context of the whole;
>
> (4) A viewpoint that gets at the heart of sociology because it sees the sociocultural system in terms of information and communication nets;
>
> (5) The study of relations rather than 'entities,' with an emphasis on process and transition probabilities as the basis for a flexible structure with many degrees of freedom;
>
> (6) An operationally definable, objective, non-anthropomorphic study of purposiveness, goal-seeking system behavior, symbolic cognitive processes, consciousness and self-awareness, and sociocultural emergence and dynamics in general (Buckley, 1967, p. 39).

The systems approach: a focus on intricate interrelationships of parts; a concern with information and communication nets; the study of relations rather than entities; the choice of organization as the central variable.

But how does the systems approach work? Surely there have been advances in the 227 years since De la Mettrie suggested that the key concept was organization.

There are a few hints.

Karl Deutsch says that only in the last fifty years have we seen the beginnings of new models that might help us. These new models, he says, have been the developments of "communications engineering":

> Modern studies of communications engineering suggest that the behavior of human organizations, peoples, and societies have important relations in common with manmade communications networks, such as servomechanisms, switchboards, and calculating machinery, as well as with the behavior of the human nervous system and the human mind. It now seems possible to analyze and describe the common patterns of behavior of self-modifying communications networks in general terms, apart from the question whether their messages are transmitted and their functions carried out by circuits of electrical current in an electronic device, by chemical and neural processes inside a living body, or by spoken, written, or other communications between individuals in an organization, group, nation, or society (1968, pp. 389-390).

Rapoport and Horvath suggest that topology, showing causal relations as directed segments, and representing complex systems as networks of interrelationships, is one important conceptual tool (1959).

In making this suggestion, they support Deutsch's focus on networks of interrelationships.

And finally, Ashby tells us that the way not to study a complex system is ...

> ... by analysis, for this process gives us only a vast number of separate parts or items of information, the results of whose interactions no one can predict. If we take such a system to pieces, we find that we cannot reassemble it ... (1956),

much as "all the king's horses and all the king's men couldn't put Humpty together again."

Clearly, then, a focus on networks of interrelationships is consistent with the systems model, which looks both at the relationships among the parts of the system and the ones between the parts of the system and the whole system. This approach will have to simultaneously look at the whole system and the parts of which it is made -- it will have to work at both levels at the same time.

This network/systems approach was first applied in social sciences in the thirties, with Moreno's sociograms (1934). In the 40 years since then, many significant advances have been made, in both the conceptual and methodological areas. As the appropriateness of the systems approach in social systems became apparent, more and more investigators turned their attentions in this direction. When this new "field" of research is examined from the perspective of 1975, however, it becomes clear that the "systems approaches" we have seen are only tentative beginnings. The conceptual basis has not yet been fully worked out, and this means that methodological difficulties will be many, since the methods depend on a clear conceptual foundation.

In Chapter Two, we will review some of the methods that have been used since 1934. We will be interested in three issues there -- what the methods do, what their conceptual bases are, and what kinds of problems they have. We will then show how most of these problems stem from a lack of clear conceptual foundations.

CHAPTER TWO

PREVIOUS METHODS OF NETWORK ANALYSIS

INTRODUCTION

In this chapter we review several methods that have been used to carry out the "systems-as-networks" approach that was introduced in the last part of the first chapter. In actual applications, the model translated "networks" in general to "communication networks." The methods reviewed here are thus all examples of an approach which might be called "communication network analysis."

Many different methods have been developed to analyze communication networks. For our purposes, they can be divided into seven categories: (1) sociograms; (2) matrix manipulation; (3) matrix multiplication; (4) factor analysis; (5) multidimensional scaling; (6) blockmodeling; and (7) a set of miscellaneous other methods. Of these methods, the sociogram, developed by Moreno (1934), was both the earliest and the most influential.

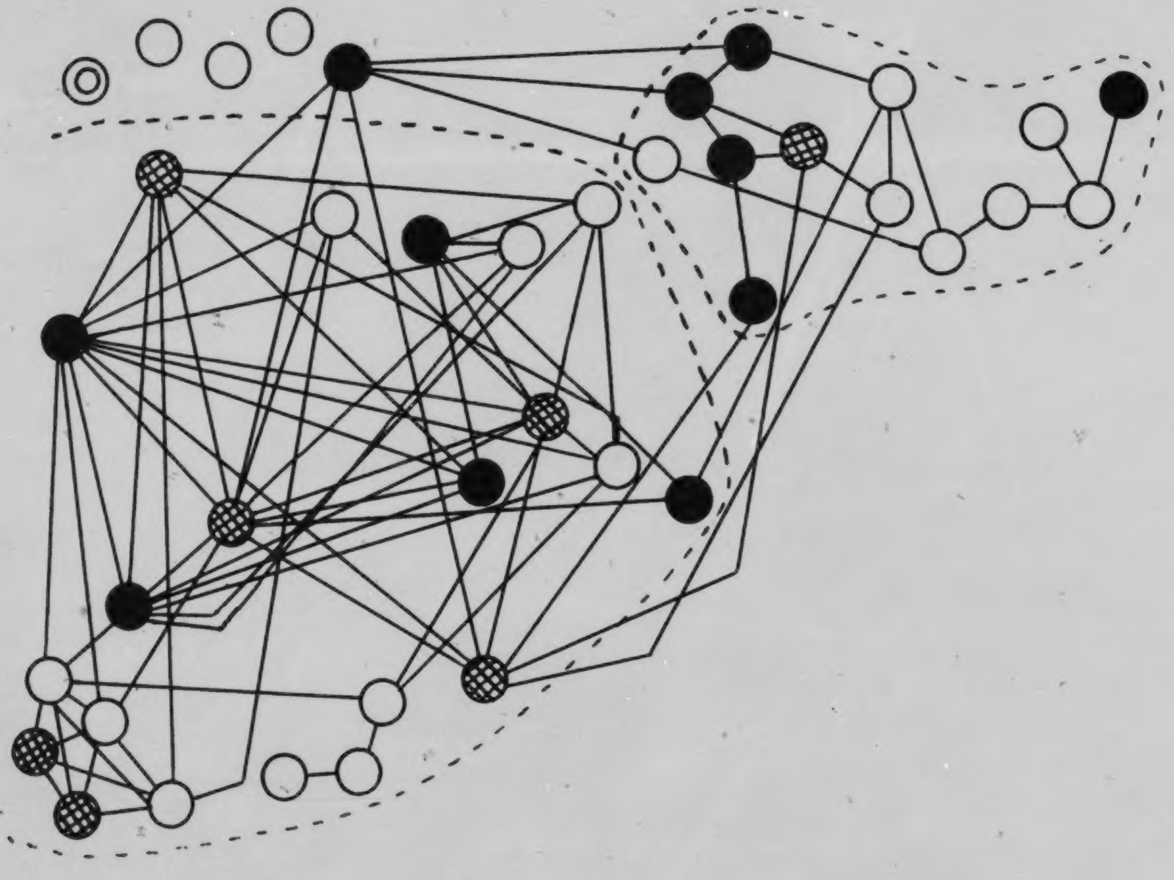
THE SOCIOGRAM

In a sociogram, individuals are represented by points, and communication relationships between individuals by lines drawn between the corresponding points. (A sociogram is illustrated in Figure 1.) By examining the sociogram for a given network, the structure of the system is determined, in terms of cliques or clusters of people who communicate primarily with each other.

The introduction of the sociogram was a major advance for the field. For the first time there was a concrete way of representing systems of interacting individuals. The process seemed quite elegant and was operationally very simple. The graphical representation of group structures provided by sociograms proved to be useful to both practitioners and theoreticians alike. For small simple systems, the sociogram seemed to be the ideal tool. Despite the good points, however, there were problems with this method.

Figure 1

A SOCIOGRAM*



---

*From Rogers, Everett M., Park, H.J., Lee, S.B., Chung, K.K., Puppa, W.S., and Doe, B.A., "Network Analysis of the Diffusion of Family Planning Innovations Over Time in Korean Villages: The Role of Mothers' Clubs," paper presented at the Population Association of America, Seattle, April 17-19, 1975.

Although Proctor and Loomis (1951) present some guidelines for the preparation of sociograms, there were no conceptually grounded analytic techniques for the analysis of systems using this approach. Instead, analysis was either by examination or by the "application" of concepts borrowed from graph theory, an abstract branch of topological mathematics (see Harary, Norman, and Cartwright, 1965; Flament, 1963; Coleman, 1964; Roby, 1968; and Lorrain and White, 1971).

Besides the lack of a solid conceptual basis, which would identify theoretically relevant goals and methods of achieving them, there were practical problems. As the size of the system increased, so did the difficulty of analysis, so that for systems having over a hundred members the method was practically worthless. It is virtually impossible to "comprehend" a sociogram for a system having, say, 200 members. This problem was further aggravated by the lack of systematic procedures that would lead to replicable results -- two independent investigators, using the same data, would seldom end up with sociograms that were even vaguely similar.

MATRIX METHODS

Due probably to two factors — the recognition of the problems connected with the sociogram, and the growing tendency to mathematize the social sciences — other approaches were developed. Without exception, these newer methods represented network data in the form of matrices, where there is a row and a column for each individual in the system. If person I communicates with person J, a "1" will be entered in row i, column j. Otherwise, the entry will be a "0". The entries do not have to be restricted to binary values, although this is a common approach (some methods require a binary matrix). The discussion of the matrix-based methods will be organized into three sections. First, we will examine each in terms of: (1) the mathematical paradigm used to represent the data; (2) how the methods go about finding "cliques"/"groups"/"clusters"/"blocks"; and (3) how they define "cliques"/"groups"/"clusters"/"blocks." Second, we will discuss the limitations or difficulties connected with each, and, finally, the relative advantages of each method. Before going into the specifics of the six classes of methods, a discussion of the "dimensions" along which we will analyze them will be appropriate.

1. Distance vs. Linkage Models

By the "mathematical paradigm used to represent the data," we mean one of two that are used -- distance models and linkage models. In the former, raw network data (who is connected to whom) are transformed to give a matrix of "similarities" (correlations) or "distances." These "distances" are defined only in terms of the mathematical operations used to compute them. They are not distances in terms of either actual physical distance or of the number of steps necessary to send a message from one node to the other. Two nodes are similar (close) if their columns (alternatively, rows) in the sociomatrix are similar (correlated). In order to transform a similarities matrix into a distance matrix, the elements in the former, which usually range from 0.0 (completely dissimilar) to 1.0 (identical), may be inverted and decreased by 1.0. Thus, a similarity of 0.0 will become $\frac{1}{0}$, or infinity. Infinity minus one is still infinity. Thus, two nodes with completely dissimilar columns (rows) in the sociomatrix will be infinitely far apart in the distance matrix. A similarity of 1.0 will become $\frac{1}{1} - 1 = 0$. Two identical-columned nodes are separated by a distance of zero.

In contrast to the distance approach, the linkage approach uses the raw sociomatrix, either as it is collected, or in a binary form obtained by deleting either weak links or non-reciprocated links, or both. Thus, a linkage matrix (sometimes called the adjacency matrix) may either be binary or continuous, with higher values referring to stronger links.

The differences between the two approaches are several:

(a) The distance approach requires a full matrix (i.e., there is a value for every pair of elements in the system), while the linkage method does not (i.e., an adjacency matrix may be largely empty, signifying the absence of relationships between many pairs of nodes).

(b) In a sense, the calculation of the distance matrix creates new information -- information about pairs of elements not directly related to one another in the raw data. At the same time, it loses much information. It is impossible to reconstruct a unique adjacency matrix, given a distance matrix, because it is not possible to tell whether or not any particular pair of nodes is linked -- only how similar their sets of links are. In fact, it is quite possible for a pair of nodes to be very "close" to one another, and yet have no direct contact.

(c) The two methods involve different scaling processes. The linkage method can use nominal, ordinal, or ratio level scaling with appropriate choice of measurement instruments. (Nominal: Who do you talk with? Ordinal: Rank the people in the system in order of frequency of contact, so that you give the person you spend the most time with a "1"; the person you spend the next largest amount of time with a "2"; and so on. Ratio: Please indicate how much time, in hours, you spend with each person in a typical month.) In contrast, the numbers obtained with a distance method may be ordinal, or possibly (doubtfully) interval, but not ratio. Infinite distances are especially troublesome.

2. The Definition of Group

All of the matrix-based methods identify some kind of "cliques"/"groups"/"clusters"/"blocks." The term "cliques" was originally applied in sociometric studies where liking or disliking relations were of interest. Because the term "clique" has affective connotations, the term "group" is used here, as it is felt to be somewhat "cleaner" conceptually. Similarly, the term "cluster" is not used because it carries less of the meaning we desire than the more descriptive "group." "Blocks" are a special kind of cluster, and will be discussed with blockmodel process. In this comparative discussion, then, the term "group" is used as a general term which subsumes all the others.

There are at least five distinct types of groups that are used in these methods of analysis. Since the definition of group varies from method to method, the particular definitions will be presented with the discussion of the methods.

3. The Method of Identifying Groups

Since all the methods identify groups (or at least detect them) they must all have a process or algorithm for doing this. Four different approaches to this process can be identified.

(a) The method of division. The entire network is divided into two parts. Each part is then divided into two more parts, and so on, until the desired "fineness" is reached.

(b) The method of agglomeration. Groups are started with a "seed" -- a node which is chosen because of some characteristic like a large number of links. Nodes are identified and added to the seed by a variety of methods, until no more nodes can be found that fit the criteria. The result is a group.

(c) The methods of trial and error. 1) Rows and columns of the adjacency matrix are simultaneously permuted to give a specific type of ordering which allows groups to be readily identified by eyeball inspection. Re-ordering rules may be informal, thus the "trial and error" name for this method. 2) In the case of factor analysis and some hierarchical clustering methods, a family of solutions is obtained and the "best" one is used; the others, having been tried, are rejected.

(d) The method loosely called overall pattern recognition. This is the method used in the analysis of sociograms. It is also used in the matrix-based methods in several different forms. The adjacency matrix (or a similarities matrix) may be reorganized in some way that groups are readily identified, either by inspection or by some process that is somehow analogous to "looking" at the whole system.*

---

*The method presented in Chapter Eight uses such a technique, overtly modelled on human perceptual processes.

THE METHODS

With this introduction in mind, we can now move to an examination of the various methods.

1. Matrix Manipulation (Forsyth and Katz, 1946; Beum and Brundage, 1950; Jacobson and Seashore, 1951; Weiss and Jacobson, 1955; Borgatta and Stolz, 1963; Coleman and MacRae, 1960)

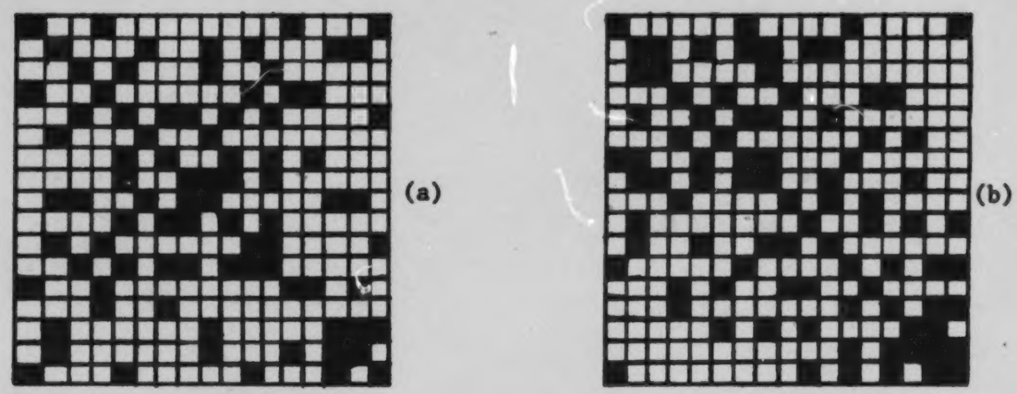
This is a linkage-based technique named after the method used to prepare the data prior to identification of groups. In this process, rows and columns of the sociomatrix are simultaneously permuted in such a way as to move as many of the non-zero entries as possible close to the main diagonal of the matrix. If there are any groups, they will be visible as clusters of non-zero entries, as shown in Figure 2. Groups are the sets of nodes whose columns and rows are in these clusters. They are loosely defined as sets of people who interact more with each other than with people not in the group. Group detection is by visual inspection, so this is a pattern recognition technique.

2. Matrix Multiplication (Festinger, 1949; Luce and Perry, 1949; Guimaraes, 1970)

This binary linkage-based method allows groups (cliques) to be detected by a process involving the raising of the raw sociomatrix to successively higher and higher powers, which allows the two-step, three-step, ... , k-step indirect paths linking individuals to be identified, as well as the original one-step links. This process allows individuals to be assessed for their integration into the system, since it allows both direct and indirect links to be counted. Although this method allows the presence of cliques or groups to be determined, the number of groups and their membership is not specified. The concept of "group" is not clearly specified here, since groups are only detected — not identified — by this method.

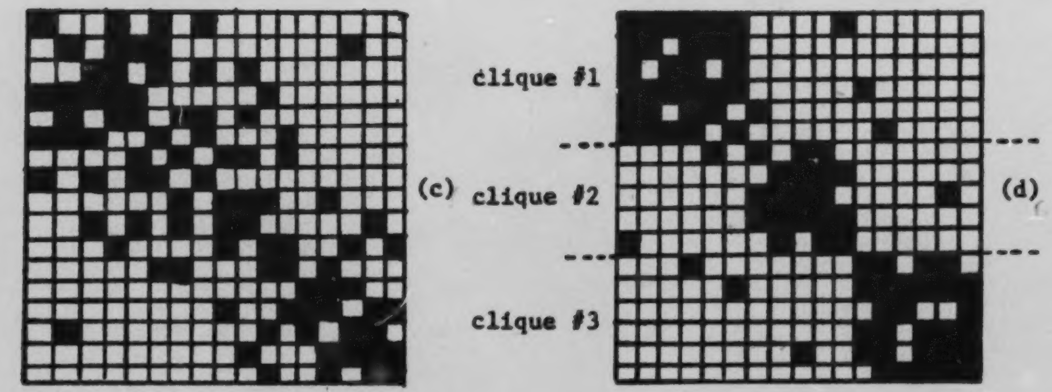
Figure 2

IMAGINARY DATA ILLUSTRATING
THE MATRIX MANIPULATION PROCEDURE*



ORIGINAL MATRIX

1st RE-ORDERED MATRIX



2nd RE-ORDERED MATRIX

3rd RE-ORDERED MATRIX

(a) shows the original adjacency matrix (■= 1,□= 0); (b), (c), and (d) are the matrices resulting from successive re-orderings of rows and columns of the original matrix. In (d) the cliques are identified as blocks of non-zero entries on the main diagonal.

---

*From Richards, 1971.

3. Factor Analysis (Holzinger and Harmon, 1941; Block and Husain, 1950; MacRae, 1960; Tyron and Bailey, 1970)

Two techniques employing factor analysis have been used -- direct factor analysis of sociometric data and factor analysis of a correlation matrix constructed from the raw sociomatrix. Both factor analysis methods construct new dimensions corresponding to variance patterns, where successive dimensions or factors account for as much of the remaining variance as is possible, using linear combinations of the variables (which are people in the network). Each factor is then a clique or group, when an appropriate arbitrary cutting point is chosen to discriminate between members and non-members (nodes loading highly or nodes not loading on a given factor). These are distance methods, where the method of clique identification is a combination of trial and error and overall pattern recognition. (The factor analysis procedure "looks" at the whole set of data and extracts patterns as factors.)

4. Multidimensional Scaling Methods (Torgerson, 1958; Shepard, 1962a,b; Kruskal, 1964; Morton, 1959; McGee, 1968; Guttman, 1968; Lingoes and Roskam, 1971)

Multidimensional scaling techniques (both metric and non-metric) attempt to identify groups from the sociomatrix by employing measures of the "distance" between the points in a "sociometric space." To use these methods it is necessary to determine first the "minimum dimensionality" of the "space" and, second, the projections of each observation onto the dimensions. Since the data must take the form of distances between persons, the dimensional axes will represent characteristics of the members of the cliques, rather than actual communication behaviors (Lankford, 1974, p. 293). Groups, then, are based on the projections of original observations onto the new dimensions.

5. Blockmodelling Methods (White, 1963, 1961 , 1970, 1974a,b; Lorrain and White, 1971; White and Breiger, 1973; Breiger, Boorman, and Arabie, 1974)

Blockmodelling methods may be either distance-based or linkage-based. We will discuss here only the linkage-based applications, because it is not clear just how the distance methods work or what they do. With blockmodelling methods, the aim is to permute simultaneously the rows and columns of either the two-dimensional binary adjacency matrix or a three dimensional matrix created by "stacking" the matrices for several different relationships, in such a way as to facilitate the identification of "blocks." A block is both a set of network nodes having similar relations to nodes in other blocks and an area in a matrix (a submatrix) identified with those nodes. Thus, all the members of Block A, for example, would be characterized as being asymmetrically rejected by members of Block B. In the other sense, that of a submatrix, there are three kinds of blocks: those having only zeroes as entries; those having some ones; and those having all ones. In blockmodelling, only the zeroblocks are of interest. These are made as large as possible by permuting the rows and columns. An ordered matrix divided into blocks is shown in Figure 3.

Blocks are unlike the other kinds of groups we have seen so far, in that "there is no implication that the members of a block cooperate or coordinate with one another. In fact, the individuals in a block need not be connected at all to one another ... " (Breiger et al., 1974, p. 10).

In a sense, then, blockmodelling is a way of categorizing the members of a system on the basis of similar interaction patterns. Thus, "it is clear that blocks need not be cliques in the standard graph-theoretical sense or any of its many sociometric generalizations" (Breiger et al., 1974, p. 10).

The matrix is re-ordered on a trial and error basis. Blocks are identified by inspection; this method is thus a pattern-recognition technique.

Figure 3

IMAGINARY DATA ILLUSTRATING BLOCKMODELS, LEAN FIT, AND ZEROBLOCKS*

(a)

$$\begin{array}{r|cccccccccc}
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
2 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
4 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
5 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
6 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
7 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\
8 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
9 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
10 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

(b)

$$\begin{array}{r ccc|cccc|ccc}
2 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
7 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
8 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
\hline
3 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
10 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
9 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Zeroblocks

(c)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

(a) a fictitious adjacency matrix
(b) the same matrix, permuted and partitioned to reveal zeroblocks
(c) a blockmodel showing relations between and within the groups created by the partition

---

*From Breiger, Ronald L., "A Blockmodel Study of a Biomedical Research Network," American Sociological Review (February 1976).

6. Miscellaneous Methods

Included in the set labelled miscellaneous is Hubbell's (1965) method which raises the sociomatrix to the p^{th} power, and sums all p matrices. This is a linkage method, with unclear goals for group definitions. It seems to fit best in the pattern-recognition class of methods.

Another method in this set is McQuitty's linkage analysis (1957), which is distance-based method, working as it does with a matrix of associations (similarities). However, this method does not function like other distance methods; it looks rather like a linkage method in the way it detects (builds) groups:

> A Linkage is defined as the largest index of association which a variable has with anh or all of the other variables. Such a linkage definition excludes overlapping of cliques and isolates. The method first joins the reciprocal pairs and then draws in all other remaining unilateral relations, which McQuitty calls clique 'cousins.' Correlation was used as an index of association.
>
> Steps in the analysis are:
> 1) Note the highest entry in each column.
> 2) Note the highest entry in the entire matrix, develop the first pair, A, B.
> 3) Read across the rows A, B, selecting any underlined entries as clique members (first cousins).
> 4) Read across rows of the first cousins, selecting underlined entries as second cousins.
> 5) Search for third and higher order cousins.
> 6) Excluding all persons already classified, repeat steps 2 thru 5 till all persons are classified (Lankford, 1974, pp. 296-7).

McQuitty's approach is clearly an agglomerative methods, with *post hoc* definitions of groups as "what the method produces."

This overview of methods is not exhaustive of all the methods that have been used. It does, however, cover the major ones.

Problems

If any single conceptual weakness is common to all the methods discussed above, it is that there was a consistent failure to specify what the analytic goals were before choosing an analytic method. Thus, investigators seem to have let the method provide the definition of group, rather than let the analytic method be shaped by the preconceived notion of what they were looking for. Therefore, in factor analysis, a group is a factor. In multidimensional scaling, groups are what the scaling routines produce.

When the goals of analysis are not clearly specified, it becomes difficult to establish guidelines that will allow different techniques to be judged for validity and accuracy. Moreover, it is difficult to justify a decision to go for a linkage-based method over a distance-based method, or to use pattern-recognition techniques rather than devisive or agglomerative techniques. If there is no clear conceptualization of the relation of the data to the original system or to the final description of the system it is impossible to know which kinds of operations on the data will be legitimate and which will not. Finally, a clear conceptualization will help by providing clues as to how to set up an algorithm that will provide fast, useful, complete results.

This fundamental weakness seems to be the ultimate cause for most of the more concrete difficulties experienced by users of the methods discussed here. We will discuss some of these problems below.

The most universally encountered limitation is one of size. Sociograms are good for up to about fifty-person networks. Matrix manipulation and multiplication are good for up to about three hundred.* The ultimate solution to this problem is to realize that in very large systems most of the possible pairs of nodes that could be related are not. Thus the storage of an entire matrix is wasteful of storage space.

The second major problem is in interpreting the results from routines originally designed for other purposes. This is a major fault of factor analysis, multidimensional scaling, and some other distance-based methods. Related to this difficulty is the fact that many methods simply do not produce useful results. Matrix multiplication, for example, can only detect cliques -- it cannot identify them. That is, it can indicate whether or not there are any cliques, but it does not tell either how many cliques there are or who their members are.

---

*Although larger computers with virtual storage seem to provide an answer to this limitation, this is not the case. While it may be possible to store a virtual matrix for a few thousand persons, swapping and paging requirements would make execution times excessively long.

The trial and error methods, as well as many of the pattern-recognition techniques, may be faulted for their lack of formal analysis rules. Again, this stems from the lack of clear conceptual foundations. It is difficult to tell someone how to accomplish a task if the task itself is not understood, let alone if it is not clear when the task has been successively accomplished.

A brief listing of the difficulties encountered with each of the methods discussed above follows.

1. Matrix Manipulation

(a) Size limitations. Not practical for over 300 person networks.

(b) The rules for some manipulation techniques require many subjective decisions and are not explicit enough to allow computerization, while other methods, which have been computerized, are very expensive to execute, requiring inordinately long computation times (see, for example, Beum and Brundage, 1950).

2. Matrix Multiplication

(a) Subject to the same size limitations as matrix manipulation.

(b) Does not identify cliques.

(c) This "awkward and very restricted" method only works for binary matrices, and handles only mutual choice (Lankford, 1974, p. 288).

3. Factor Analysis

(a) Suffers from similar size limitations as the earlier matrix methods.

(b) The two computational stages of this method, factoring and rotation, consume large amounts of computer time relative to other methods.

(c) The assignment of nodes to groups is somewhat arbitrary since the investigator must choose the size of the factor loading to use as a cutting point, as well as the number of factors to rotate.

(d) The results obtained from factor analysis are difficult to interpret and use as descriptions of actual communication behaviors.

4. Multidimensional Scaling Methods

(a) These methods suffer from the same size limitations as factor analysis, and use even more computer time to execute.

(b) The final results are not at all close to the original data, due to the many complex transformations that are made on the data. Indeed, Lankford (1974, p. 301) criticizes this method as being "roundabout."

(c) The determination of groups is not a clearcut procedure.

(d) As in factor analysis, the results are difficult to utilize and interpret. Lankford says that "multidimensional scaling has proved to be a very inefficient method for clique identification both empirically and logically" (Lankford, 1974, p. 301).

5. Blockmodelling Methods

(a) Suffers from size limitations similar to those of matrix manipulation.

(b) There are no clearcut procedures for putting the adjacency matrix into the form needed to construct a blockmodel. The procedures used by the originators of this method are described as "trial and error" (Breiger et al., 1974, p. 11).

(c) The relation of blocks to actual groups or cliques is unclear, and may vary from case to case. This method seems to be more of an individual level of analysis, although it does function in the context of the larger system.

6. Miscellaneous Methods (The set of other methods, including Hubbell's and McQuitty's methods)

(a) Suffer from similar size limitations as matrix multiplication methods.

(b) Seem to not work reliably, sometimes producing breakdowns into groups and sometimes not (Lankford, 1974, pp. 295-6, 296-7).

(c) Fail to clearly state the goals of analysis, resulting in ambiguous or vague relations between results and raw data.

Relative Advantages

The new methods promised many benefits and were able to deliver on some of these promises. The sociogram provided a way of graphically representing network structures. The matrix techniques provided a mathematical format for the data, and offered investigators a more highly structured way of representing complex situations. The power of matrix representations of other kinds of data in other kinds of situations promised to be fruitful also in the social sciences. The similarity of sociograms and formal graphs (in the topological sense) led to the application of graph theoretical concepts to communication networks. One fallout of the use of graph theory is the matrix representation of the network, since matrix representations of formal graphs had proved very useful.

ON MATRIX REPRESENTATIONS

The matrix is an especially powerful form of mathematical notation. Using matrices, one can speak of sets of relations and operations that would be prohibitively cumbersome with less sophisticated notations. The "condensation of meaning" (Brown, 1969, p. 81) into matrix operators is what makes them so powerful and elegant. It was the hope for an extension of this same power and elegance that led to the use of matrices in sociometric and network situations.

Compared to complicated and unruly sociograms, matrices are very well-behaved indeed. In a matrix the numbers are arranged into neat rows and columns, and even the order of the rows and columns carries meaning. It is easy to locate any particular element in a matrix, simply by giving its subscripts.

This ease of location and manipulation was probably one reason matrix representations were favored in computerized programs for network analysis. Most of the popular scientific computer languages that are good for numerical work, like FORTRAN or ALGOL, are arranged around array structures. An array is a matrix of one, two, three, or more dimensions. Elements in arrays are referenced by means of subscripts, which may be orderly varied by simple computer statements like DO-loops.

Thus, matrix representations became popular early, and became more deeply entrenched with the advent of computers and computer programs. This popularity can be seen with a quick analysis of the techniques available in the field -- with the exception of the sociogram and the method presented in Chapters Seven and Eight, every one on which information was available used matrix representations.

Although the matrix representations did offer much in the way of utility and theoretical promise over the simple sociogram, and although they did allow a host of different kinds of analytic techniques to be tried, they may have hindered progress in the field. They seem to be so ideal and "natural" that there was little tendency to look for other ways of representing networks and performing analyses.

The use of matrix representations was symptomatic of a deeper problem, however. This deeper problem has been discussed already -- it is the lack of clear conceptual foundations. In one way or another, all of the computational difficulties listed above for the various methods can be traced to this one fundamental lack. It is this antecedent problem that we focus on in this thesis. The position taken here is that a better understanding of networks will lead to better ways of studying them, which will lead not only to better access, but also to more powerful theory about how networks work.

EXTANT DATA

In spite of the difficulties mentioned above, massive amounts of empirical data have been gathered on communication networks. These data come from a large number of studies, most of which were experimental or quasi-experimental laboratory studies. The studies themselves will not be individually reviewed here, as there are already several extensive reviews in print (see, for example, Collins and Raven, 1969). Two considerations, however, preclude the use of most of this empirical information. The first is that almost all of these investigations were done on systems having three, four, or five members. Two serious problems stem from this choice of small sizes.

(a) There is neither conceptual nor empirical agreement whether generalization is possible across these three group sizes. Investigators not only get different results, but they cannot agree on how to define certain situations across group sizes (Collins and Raven, 1969, p. 147).

(b) Findings based on systems having five or less members in a laboratory situation cannot possibly be extended to real functioning systems having several hundreds or even thousands of members. Five-person groups are simply too small to allow the kinds of things commonly observed in large systems, such as hierarchical organization to occur. In addition, the systems used in this work have typically had lifetimes of less than two hours. This temporariness can be contrasted with most real-world systems, which have been evolving and growing for years. Certainly, a group of several hundred people who have been working together for many years is hardly comparable to a group of five students who have been working on a puzzle for a few minutes, or even to a small group experiment lasting three days in a simulated fall-out shelter.

Second, according to Collins and Raven (1969, p. 147), an unfortunate state of affairs is "prevalent throughout the entire communication-net literature." They say, "It is almost impossible to make a simple generalization about any variable without finding at least one study to contradict the generalization," and go on to suggest that "in the light of the massive confusion present in the literature, the time may well have come to tighten our statistical belts." They feel that the unreliability of the findings is due to the "liberalness with which most investigators have treated the traditional .05 level." They seem to be taking an approach which says, "we're sure there's something there; we just haven't looked close enough."

CONCLUSION

A major point of this chapter is that we've looked neither in the right places nor in the right ways. This point will be extended in the next part, where we will expand the conceptual paradigm upon which the methods we discussed in this chapter were based, by examining the entire problem of complex systems from what might be called a "naive" perspective. That is, we will attempt to approach the situation with as few preconceived notions as possible so that, perhaps, our approach will be fresher and, hopefully, more fruitful.

In Part Two, then, we will be looking for answers to three questions:

(1) What are we looking at?, i.e., what are complex communication systems? How are these systems organized? How do they come to function as coordinated units, when none of their characteristic properties are seen in their component parts? How are the parts related to one another and to the whole system?

(2) How do we look?, i.e., how, in general, are the processes of observation and description carried out? How does the relationship between the observer and the system being observed influence the processes of observation and description? What happens when the system being observed is complex and multi-leveled?

(3) What do we look for? When we combine an understanding of some basic systems principles with an awareness of

> how the processes of observation and description work, how can we structure the investigation process so as to gain as much useful information about complex systems as possible? In other words, what will we be looking for when we study complex systems?

The answers to these questions will provide the basis for a coherent systems approach to the study of complex systems, viewed as networks of interrelated parts. In the last sections of Chapter Five, we introduce a procedure for performing this type of analysis — GASSICS — a General Algorithm for Studying Structuring In Complex Systems.

In Part Three, we will discuss an initial operationalization of the first steps of the GASSICS procedure. Even though the procedures presented there are very complex, they only take us through the earliest stages of a complete systems analysis. Later stages will have to wait until much more data than we presently have is collected and analyzed; until much more theoretical work is completed; until more sophisticated computer programs are written to perform the necessary analyses. Since this is the work of many years, only the first stages will be included there.

In the fourth part, we will present an example of this method. We will perform an in-depth analysis of the structure of a large-scale communication system. This analysis illustrates the use of some of the procedures described in earlier chapters.

Finally, in Chapter Eleven we will look toward the future to see where the logical next steps should be. Here we will be trying both to anticipate and direct further work in this area. While we cannot yet tell what the results of ongoing research will be, we are aware of some limitations to presently available methods. An analysis of these limitations will suggest specific areas that need more work if we are to continue to make rapid advances in this field.

PART TWO

CONCEPTUALIZATION

In the first Part we discussed the evolution of the basic paradigms used to structure approaches to the study of social systems. The discussion of conceptual approaches in Chapter One was followed in Chapter Two by a description of the operational techniques that are used to perform the actual analysis.

While most of the methods discussed in Chapter Two were far superior to the earlier organismic or mechanistic approaches, none was ideal. Even the best of the methods, in terms of efficiency or appropriateness, had problems. We identified one problem as common to all the earlier methods — a lack of clear conceptual foundations. We showed in Chapter Two how this most fundamental weakness led to a plethora of problems — vague analytic goals, unclear standards to use in judging analytic methods, inefficient techniques, unclear definitions, and a large set of points over which no general agreement has been reached.

Examining the conceptual void more closely, we pointed out the necessity to obtain a better understanding of the form of the system we want to know and of the processes by which we gain that knowledge — the processes of observation and description.

In the chapters of this Part we start out at the beginning — with a re-conceptualization of the concepts fundamental to the study of systems. We begin in Chapter Three with an analysis of the form of systems. Form is the ultimate basic. The discussion moves into other areas then, with the introduction of concepts one step away from form — constraint, interaction, and elements. This is all still at the level of simple systems. With the introduction of emergent properties — properties due to the interaction of elements, we move into the realm of complex multiple-leveled systems. The same basic concepts are discussed in this new context, and the implications are explored. The discussion of interactions between complex systems ends with the concept of information, which appears as a shorthand way of describing a whole class of interactions.

Chapter Four examines the processes by which we gain access to form -- the processes of observation and description. The development of this chapter is parallel to the development of Chapter Three. We begin with an examination of the concept of description. This is seen as a process of drawing distinctions, based on differences perceived in the form. The form of descriptions is examined and related to the process of observation, which allows the differences upon which the description's distinctions are based to be "seen." This is all done at the level of simple systems. Again, with the introduction of emergent properties and complex systems, we move into a new realm -- that of alternate descriptions.

The processes of observation and description in complex multi-leveled systems are logically very complex. We explore the ramifications on the basic ideas in this complicated situation in the context provided by the analysis of Chapter Three. Since all observation processes are interaction processes, and since the observer is himself a system, the observation process is limited as are all interaction processes between complex systems. These limits, together with the ones inherent in the process of description, place limits on the way complex systems may be approached. The chapter ends with a discussion of these limits in the context of alternative descriptions.

Chapter Five picks up at this point with an examination of structure as the appearance of form. If form is what is, structure is what the observer sees. Again, we go back to basics, this time to constraint and form. We develop the relation between form and its appearance; and here the centrality of constraint is unique. The kinds of constraint that can be observed will determine what can be said about the system being studied. "Properties" are based on constraints, and thus contraints determine properties. The search for understanding is thus a search for constraints, which appear to observers as different kinds of structuring, in both space and time.

When we pull on the common thread running through the discussion in Chapters Three, Four, and Five, we get a general approach for studying complex systems. The starting point is structuring -- the appearance of constraint. We call the procedure that falls out of the discussion a General Algorithm for Studying Structuring in Complex Systems -- GASSICS. The introduction of this procedure ends the last chapter of Part Two.

CHAPTER THREE
THE FORM OF SYSTEMS

INTRODUCTION

We have seen so far how people have tried in the past to analyze the networks of interrelationships among the members of social systems. We identified the main problem with these approaches as a conceptual one. Indeed, there was remarkably little in the way of solid conceptual foundation for this approach. We can look to the field of "systems theory" to find just the opposite situation. Here is a "theory" — a very successful one, if judged in terms of popularity — with no methodology (at least, not as it has been used in the social sciences). It is difficult to think of a traditional scientific discipline in which this "approach" has not been "applied." Biology, psychology, sociology, anthropology, medicine, psychiatry, political science, and economics are only a few of the areas that have used the concepts of the "systems approach." In spite of the very wide set of areas using the concepts of systems, there are serious problems, both with the "theory" of systems theory, and with the methodology available to investigators who wish to apply the concepts to their own areas of interest.

We are interested here in a particular kind of system — the human communication system, in which the basic process of control and coordination of the parts is communication. Now, a concept central to the study of communication is the one of information. All communication processes are at their core information processes. Information is what is communicated. Without information there is no communication. The study of communication processes is thus the study of information processes.

Indeed, information is such a central concept that a whole "theory" has been built around it, called "information theory." Here again, however, there are problems. The "theory" of information is incomplete and ambiguous. Different people have different meanings for "information." There is no "general theory of information/information processing." Most discussions of information processing seem to be either case studies, precluding generalizations to other cases, or else at such a high level of generality that they say nothing about everything. The central concept of what is usually called "information theory" (Shannon and Weaver, 1949) — information -- is a strange concept indeed. The "theory" provides elegant mathematical tools which can be used to measure the amount of information, but "Only the amount ... is measured — the amount does not specify the content, value, truthfulness, exclusiveness, history, or purpose of the information" (Miller, 1968, p. 123). We can only say how much there is, and even this only in a highly restricted context. The theory does not clearly say exactly what this "information" is, in terms of what it does when it gets where it is going, or how it got to be the way it was. Because the theory is not well developed, a rigorous methodology for studying information processing systems remains to be seen.

A few attempts to apply information theoretic concepts to human communication situations have been made (see, for example, Garner, 1962 ; and Danowski, 1974). In many of these the "application" consists merely of the use of the information statistic as a measure of variance at the nominal level: "How much uncertainty is there in this situation, where we have X alternatives with probabilities of P_1, P_2, ... P_X?"

The use of "information theory" as an approach to gain a better understanding of the communication process as an information process, in (scientific) theoretical terms, has not been seen. This is to be expected, given both the traditional orientation for communication research and the a-theoretical nature of "information theory." At this point many would argue that the advent of "systems theory" will change this. We agree that the systems approach is most promising, but this is an area that is almost as young as "information theory"; and it suffers from many of the same kinds of problems as the latter.

Although more people have written about systems or systems theory than about information theory, the key concepts have yet to be drawn together in a way that provides a consistent systems methodology. Many writers use the vocabulary of systems theory in their work, and are able to provide more elegant descriptions of complex phenomena than could be had with more traditional analytic approaches. Indeed, concepts like goal-directedness are hard to describe within the traditional linear framework of mechanistic science, which has been so very successful in the physical sciences. The vocabulary alone provides concepts that banish mysterious vitalistic forces like entelechy and replace them with respectable cybernetic control mechanisms and emergent properties which have clear foundations in the parts of which the systems in which they are seen are made.

The vocabulary alone is not enough, however. With the new concepts we need new ways of looking at things -- we need a new methodology -- new tools. Nowhere is this more evident than the social sciences. People say "We must take a systems approach," but this seems to mean "We must speak the language of systems." There are no methods that were developed in the context of systems theory, no techniques whose assumptions are consistent with the assumptions of systems theory. Up to now, the best we have been able to do with these systems concepts, says Deutsch (1968, p. 390), is treat them qualitatively, by recognition or description. We have no quantitative methods. Why is this the case?

One reason, perhaps, is because the mechanics of the processes of observation and description, which include the measurement and analysis of data, have not been considered to be as important in the systems situation as they really are. Our preliminary formulation suggests that these processes are much more complex than has been recognized. This is especially the case for information processing systems.

In this chapter we examine the logical form of systems and identify some crucial concepts that are basic to the paradigm we are working from. A key concept is the concept of interaction. The most important, in terms of the impact it will have on everything we do with systems, is the concept of levels. Even more fundamental than these most basic systems concepts, however, is the concept of distinction, as explicated by Brown in the Laws of Form (1969).

Very simply, a distinction is a division into exhaustive and mutually exclusive parts, on the basis of some difference which is perceived by the observer. Distinctions are arbitrary, and are drawn by the observers of systems. A distinction is not drawn unless the contents are seen to differ in value, and there can be no value without motives. Thus the motives of the observer will determine his values, and therefore the distinctions he draws.

Our motive is to understand complex information-processing systems in general and human communication networks in particular.

Although many of the distinctions we draw in the formulation of our model may seem to be highly abstract and perhaps even a bit metaphysical, they all have major implications for the later, more concrete, aspects of the model and the uses to which it can be put. Specifically, they will determine the form of the methodology we are proposing and suggest the directions in which we will move with the theory based on the new paradigm.

We start with a look at systems — the context in which "information" makes sense. The kinds of system we will be interested in are all multi-leveled. This means that the system as a whole is made of parts, which themselves are made of smaller, simpler parts, and so on. In addition, the systems are complex — taken as wholes, they show properties not seen in their parts taken independently.

The suggestion that information-processing systems must be complex (if a system is complex, it must also be multi-leveled) stems from the notion that "information" acts on the system by influencing small parts of it, rather than by acting on the system in its entirety. (It turns out that in systems which must either act as units or else not at all, the capacity for "information," as we will define it here, is exactly zero.) The requirement of complexity is associated with the notion of what are called "emergent" properties — properties due to the way the parts of the system interact. A system with strictly additive properties, where the characteristics of the system as a whole are directly (additively) related to the characteristics of the parts of which it is made, does not appear to process "information."

At this point it should be sufficient to note that the restriction to complex systems is not counter-intuitive or illogical. Included in this category of systems are units such as individual people, small organizations, computerized information processing/retrieval systems, computer networks, and even large societies. These are exactly the kinds of systems we are interested in.

We begin with some basic concepts, and from these we move to a very simple and orderly kind of model. We work from this simple model to one that will perhaps be sufficiently complex to help us see how we can begin to account for some of the more interesting aspects of "information processing" systems in general and human communication networks in particular.

1. Interaction

The conceptual basis for our paradigm is exceedingly simple: "Macroscopic objects [systems] are complex structures of [microscopic] ones. The properties and relations of the former therefore occur under conditions that can be formulated in terms of the arrangements and interactions of the latter" (Nagel, 1961, p. 312). Three ideas are brought together here: system, arrangement, and interaction. A system is a complex structure of smaller parts. These parts are arranged in some particular way, and they interact with each other. We start with the process of interaction, which is logically the most basic of the three concepts.

We say one entity interacts with another when there is some co-variation of some properties or state variables of the entities involved. In other words, there are fewer degrees of freedom in the interacting set of entities than in the set of entities taken independently. There is some constraint of the freedoms of the entities, which results from/indicates/is the interaction. Ashby says that the constraint is the relation -- it "occurs when the variety that exists under one condition [independence] is less than the variety that exists under another [interaction]" (1961, p. 130). Thus, when a set of entities is juxtaposed in some space (spatial, temporal, or conceptual) in such a way that the degrees of freedom of the entities, when they are in the set, is less than the sum of the degrees of freedom of the entities taken independently, the entities are said to be in a state of interaction.

For example, say we have two blocks of wood on a tabletop. Now, each piece can rotate, or it can slide in any direction on the table. If we count rotation as one degree of freedom, vertical motion as another, and horizontal motion as another, we see that each piece has three degrees of freedom of motion. (We have ignored velocity for the sake of simplicity.) In the set of two pieces, all possible combinations are allowed, and there are thus six (3+3) degrees of freedom. If we put the pieces together so they interlock, they can only move as a unit. The two pieces together now have three degrees of freedom. The interaction is evidenced by the reduction in the total degrees of freedom of the system. In this case, there is co-variation of movement of the pieces in the interacting set. That is, if one moves vertically, so does the other.

2. Levels

A second very important basic concept is the one of levels. Several points must be made here.

(a) As a concept of distinction. First, and most importantly, is the idea that the concept of levels is a concept of distinction/description -- it is artificial: imposed by the person describing the system. Since it is a concept of distinction, it should be defined in such a way that the distinctions we make by using the terms are both useful and consistent. It is important to be aware of this point, because the way we see things initially is sure to influence the things we work with later -- the things we single out as important, the way we perceive "things" or "dimensions" or "units," for example.

(b) Multiple levels. The indication that there is one level implies that there must be/could be other levels. Otherwise it would be pointless to distinguish between levels, and thus pointless to use the term. Conversely, there can only be levels if there is a difference between them — if a basis for distinguishing what is at one level from that which is at another can consistently be followed.

(c) Hierarchical levels. The term "level" is used to distinguish between classes in complex systems, which are made of parts, which are made of smaller parts, and so on (Pattee, 1973). The hierarchical nature of these systems leads naturally to cases in which the elements at one level may contain or be made of elements at another level. This fact implies that a different kind of distinction is being made here than is usually made when one distinguishes between kinds of entities: that the present distinction depends, to an extent, on the relationships between the units at one level and the units contained by or comprising those units. We might clarify this point by differentiating between "vertical" and "horizontal" distinctions as follows: A horizontal distinction is made between mutually exclusive entities. The entities so distinguished must be separate and identifiable. A vertical distinction, however, is made between a class of some kind and a member of that class. These are special kinds of classes, and this idea will be clarified in a little while.

(d) Additive and emergent properties. We mentioned earlier the ideas of "additive" and "emergent" properties. The distinction between these two kinds of properties is crucial here. Say we have a set of elements of some kind. Each of these elements has several properties. We may combine a number of these elements in some way to form larger units. Now, the properties of these larger units can only be due to two things: 1) the summation of the properties of the original elements that were combined to make the unit; or 2) the result of the interactions between the elements that were combined to make the unit. Since the first type of property is obtained simply by adding up all the individual elements' characteristic values for these properties, these new properties are called "additive" properties. It is not necessary to know how the elements are related to each other in the larger unit to understand the properties or characteristics of the larger unit. An example of this kind of additive property is mass. The mass of a large set of objects is simply the sum of the masses of the individual objects.

Emergent properties are very different. They become apparent only when the original units are combined into sets and allowed to interact, and thus are called "emergent" properties of the larger sets or units. Two points should be evident here: first, the emergent properties "belong" to the larger set of interacting elements, and cannot in any way be identified with the elements making up the set; second, in order to understand how these properties come into existence and how they work, it is necessary to examine the way in which the elements interact with each other in the context of the larger unit of which they are parts. (Nagel has a different meaning for "emergence" than we do. He requires that a property arising from the combination of elements into a unit be unpredictable, given both full knowledge of the properties of the elements and full knowledge of the relationships among them in the larger unit, before he will call it an emergent property. In his definition, an emergent property cannot be understood by examining the system in which it is observed. It is a new property, it is emergent, it cannot be explained. We disagree with his approach on this point. [Nagel, 1961].) An example of an emergent property would be the chemical properties of any element. These properties are determined by the way the protons, neutrons, and electrons are arranged in the atoms of the element. The properties of oxygen, for example, are not seen in any proton, neutron, or electron. But put the right combination of them together and look at the combination instead of the parts of which it is made, and you have oxygen, with a whole new set of properties, even though the protons, neutrons, and electrons are still there, and still protons, neutrons, and electrons.

Our definition of levels makes use of these concepts: If, as a result of allowing a set of entities to interact, there are properties observed that are different from the properties of the entities, taken independently, the set of entities is said to be at a higher level than the entities making up the set. In other words, if there are emergent properties in a set of elements, the set is at a higher level than the elements of which it is made. The emergent properties form the basis for the distinction that is meant by "level." If there are no such emergent properties, no distinction can be made, and the set of elements is not at a higher level than the elements themselves. If this is the case, it is only necessary to understand the individual elements making up the set if one wishes to understand the set. On the other hand, when studying an interactive set with emergent properties, it is necessary also to study the way in which the elements interact with each other when they are in the set.

(f) Levels vs. types. Another distinction: Levels are not the same as Types, as in Russell's Theory of Types. In the latter, any set or class is of a higher type than the members of the set or class. In the present definition of levels, however, the set is of a higher level if and only if it shows properties different from those observed in the elements taken independently. In other words, levels are closely related to interaction: "the more the members of the class interact ... rather than coexist, the more does the superordinate group [the class] become a true individual rather than a collection of ... individuals" (Gerard, 1968, p. 53). In the latter case, no new level is observed; while in the former, a new one is, and is based on the distinction allowed by the emergent properties — the properties resulting from the interaction of the elements.

(g) A warning. An important notion to keep in mind is the arbitrary nature of any distinction-called-level. There are probably an infinite number of ways in which a system can be described that preserve the essential distinctions demanded by the definition of level. The description that focuses on significant (useful, valid, consistent) emergent properties will be likely to be arranged into descriptively significant (useful, valid, consistent) levels. What is useful will depend on the goals of the investigator of the system. What is valid will depend on the perceptual and analytic abilities of the investigator; and what is consistent will depend on the proper choice of membership criteria -- if these are ambiguous, the description will be ambiguous, and distinctions will be neither continent nor consistent.

3. Interactions Between Systems: The Form/Logic of Information

In order for there to be interaction between two entities, we said that there had to be some kind of covariation of the entities. An important class of interactions includes all the cases where the entities involved in the relationship are themselves complex systems. The existence of multiple levels in the systems involved is the distinguishing characteristic of these interactions.

(a) Interactions and levels. What does it mean to say that twp complex systems are in a state of interaction? It means that there is some covariation of the systems involved. Very straightforward. But complex systems have different kinds of properties. Some are additive while others are emergent. Which ones are covarying, and what is the difference? When I interact with the computer terminal on which I am writing this chapter, the interaction involves emergent properties. My conscious thoughts (an emergent phenomenon) usually (sometimes I make errors) covary with the letters that appear on the paper (an emergent property of the terminal, determined by the way its parts are related to each other, and by the way the computer was programmed to behave). But how do I get any single letter to appear on the page? By striking a key with my finger. Isn't the interaction between my finger and the key similar to the one between two billiard balls on a collision course? Indeed, when viewed at this level, the interaction involves mainly additive properties, even though the same interaction, when viewed from the point of view of the whole system, involves emergent phenomena.

These two kinds of interactions are fundamentally different, and this difference must be recognized when studying complex systems.

(b) Matter/energy interactions vs. information interactions. The two types of interaction introduced above can be more explicitly distinguished as follows: The first type involves the covariation of additive properties of the two systems. Either the systems interact directly, establishing the covariation of high level properties, or they interact at a lower level, such that an indirect covariation of the additive properties at the higher levels of the systems is established. These interactions will all involve matter and/or energy. They are clearly understood by applying basic physical methods of analysis. An example of this type of interaction is seen when one billiard ball strikes another. A second interaction of this type occurs when I am struck by a milk truck when I attempt to cross the street. The momentum of the truck is partially transferred directly to my body in such a way as to keep the total momentum of the system, including both the truck and my body, at a constant. In these cases, the first law of thermodynamics is obeyed. The total amount of matter and energy remains constant throughout the interaction. These interactions are called matter/energy interactions, because they involve exchanges of matter and/or energy.

The second type of interaction involves the establishment of covariation between emergent properties of the systems. Since emergent properties are due to both the properties of the elements at lower levels and the way they are related to each other, interactions involving emergent properties must involve both the elements at lower levels and the relations between them. The interactions with the lower elements may be matter/energy interactions. The effect of the inclusion of the relationships between the elements as we move from low to high levels is to make the actual form of the covariation at the higher level different in a crucial way from the one at the lower level. Specifically, the actual form of the variation at the high level will be determined by a combination of the form of the interaction at the lower level and the form of the interactions between the parts at the lower level.

This is really quite simple, and can be clearly illustrated with the example of the computer terminal. The covariation at the high level involves a matching of my conscious thoughts with the letters that appear on the page. This covariation is established by my striking of the keys of the terminal. When I strike a particular key, the thing that causes the particular letter to appear is the way the parts of the terminal are arranged. If I change the connections between the keys and the letters, different letters will appear when I strike the same keys. If I change the program governing the way the text editor works, some letters may even cause the entire text to be erased. Due to the primary interaction at the matter/energy level, combined with the particular organization of the machine, the secondary covariation involving the particular letters that I want is established at the higher level of the system.

These secondary covariations -- covariations at higher levels involving emergent properties, and due to the primary matter/energy covariations at lower levels -- have some unusual characteristics. One is that the conservation law does not have to be obeyed in these interactions, although at every step along the way it is obeyed. The reason for this is because the system may have its own source of energy, which is included in the process because of the way the parts are arranged. This is the case with the computer terminal, which is electric: the letters strike the paper with much more force than my fingers strike the keys. It doesn't matter if I press the keys harder; the letters strike the paper with the same force. A second unusual characteristic of these secondary covariations is that they are arbitrary, in the sense that they depend on the particular organization of the parts of the system, and not only on the form of the primary interaction. Because of these characteristics, interactions involving secondary covariations of emergent properties are called "information" interactions.

(c) Information. This choice of names is consistent with other models of information. Deutsch (1968) writes that "Unlike 'matter' and 'energy,' 'information' -- that is, the pattern that can be abstracted ... [from these exchanges] -- is not subject to the laws of conservation. It can be created and annihilated" (p. 392). Another definition of information requires that a "marker" bearing a pattern be exchanged in information exchanges. The marker is not the information, but bears the pattern, which is the information. Berlo called information "patterned matter/energy." In an information interaction, the units at the primary level of interaction, that is, the units involved in the matter/energy part of the interaction, would be the markers bearing patterns which are "information" to the system as a whole. As far as these units are concerned, the interaction is strictly matter/energy, while, in the larger world of the system which is connected to the elements by emergent phenomena, the interaction may be an information exchange. For example, a light pattern impinges on your retina. This is a purely matter/energy interaction, as far as the cells of your retina are concerned. If the pattern happens to form letters and words which make a sentence, the larger system of the person that is you may "receive some information" from the interaction. Whether or not this is the case will be impossible to tell by looking at your retina; it is necessary to look at the way the retina fits in with the rest of the system, because this is the context in which it functions, and the context in which the interaction must be interpreted.

This framework suggests a different interpretation of "information" than the purely quantitative one of Shannon and Weaver (1949). Their framework was one in which only the relation between the observer and the pattern was important. The framework here is one in which the parts of the system at one level are related to the parts of the system at other levels. In this model, the relative positions, in terms of levels, together with the specific nature of the relationships of the parts at one level to the parts of other levels, are important. The observer is describing what he sees (or thinks he sees) of the cross-level relationships, but he knows the statements he makes are dependent on the relationships between himself and the system.

One point suggested by this model (perhaps implied in its definition) is that information interactions can only take place in systems of complexity sufficient to allow the interactions to take place at some level lower than the level of the systems themselves. This is because information is represented as patterned sequences, which are conveyed through the markers which are exchanged in information transactions. The markers must be at a level lower than the system as a whole, or else the system would be capable of receiving zero information, since it would have to covary identically with the marker if it were at the same level. On the other hand, matter/energy interactions may take place at the same level as the whole system, and thus complexity is not required for those interactions.

Another point is suggested when this model is examined in the light of Ashby's Law of Requisite Variety, which states that uncertainty in one system can only be compensated for by another system to the extent to which the compensating system has sufficient variety. In other words, "variety can only be destroyed by variety" (Ashby, p. 135). Applied to systems in the context of our model, the law says that if a complex system is encountered by a simpler one, the covariation established as a result of their interaction can be at a level that is no higher than the highest level of the simpler system. A simple system cannot interact with a complex one because there is not enough variety in the simple one to accommodate that in the complex one. Therefore, if forced to encounter units of greater complexity, elements will interact with parts of those units, rather than with the units as wholes.

A general interpretation of this law for matter/energy systems reads much like Russell's Theory of Types (1971), which would say that elements can interact only with other elements of the same complexity. If forced to encounter a unit which is a set of elements of similar complexity, the first element will not interact with the set as a whole; rather, it will interact with a member of the set. This restriction seems to hold for matter/energy systems, but fails completely when information interactions are allowed. The nature of this failure and the reason it is of interest are due to the way information is related to matter/energy. Recall the distinctions made earlier: 1) information exchanges demand complex systems, because there must be levels below the level of the system, on which the markers bearing the patterns which are the information are exchanged, in order for there to be information exchanges; 2) the effects of the exchange of matter/energy patterns at low levels must be conveyed to higher levels through emergent properties created by the interaction of the lower parts of the system. This implies that, in any information exchange, a thorough understanding of one of the interacting systems is not sufficient to predict the effects of the exchange on the other system, unless the two interacting systems are identical. This is a direct result of the arbitrariness created by the emergent properties demanded by the definition of information. It also implies that the pattern in the exchanged markers is free of information unless viewed in the context of the system which processes that pattern. There is no meaning in the pattern. It is impossible to measure the information content of the pattern, as long as the nature of the system which must process that pattern is unspecified.

The interdependence of what we have been calling "information" and the system in which it has existence/can be observed/can be measured/makes a difference cannot be underemphasized. Information was defined as something that required some kind of patterned matter/energy. It has also been defined as a "difference that makes a difference." This alternative definition implies that a context is needed before there can be information. This context is the system in which the pattern "makes a difference." Therefore, any pattern may become information, if there is a system that can process it correctly. This is not to say that, for a given pattern and two processing systems, A and B, the "information" "extracted" by system A will be the same as that extracted by system B. Indeed, system A may be a subset of system B, so that in system B there is an extra level of emergence transforming the output of system A, which results in a higher order of "information" from the same input.

Depending on the way the parts of a system are interrelated at each of the many levels there may be in the system, a given matter/energy input at a low level may have differing effects on the system. Some inputs may remain at the level of matter/energy, influencing higher levels only through additive properties; while others may be processed as information, influencing higher levels through emergent properties. Others may do both.

CONCLUSION

In this chapter we have explored the form of complex systems. Although most of the discussion was concerned with the establishment of a precise vocabulary through the explication of some crucial concepts, we did point out some general properties of systems. The importance of some of these properties -- especially the ones restricting interactions between systems of different types -- will be brought out in the next chapter, when we look at the processes of observation and description.

CHAPTER FOUR

THE PROCESSES OF DESCRIPTION AND OBSERVATION: ACCESS TO FORM

In the first chapter of this Part, we examined the form of systems. We chose to start there because we felt uneasy about the way people have looked at systems in the past, and because we felt that our dissatisfaction was caused largely by an apparent mismatch between systems and the way we perceive, observe, describe, and (mis)understand them -- in short, because we don't know how to "look" at them. In this section we discuss the process by which investigators gain access to the form of systems: the processes of description and observation.

We contend that a major goal of all science is to provide accurate, useful descriptions of phenomena in the world. Descriptions allow us to understand, to explain, to predict. More and more, we are faced with problems involving complex systems: the economy, congress, towns, corporations, school systems, the armed forces, the United Nations. Clearly, we would benefit greatly if we could understand, explain, or predict these systems. If we could provide ourselves with the right kinds of descriptions of these systems, we could do these things.

We contend, therefore, that an understanding of the process by which we provide descriptions is a fundamental necessity. But first, we must understand what a description is -- what is does, how it is constructed, how it is related to the form of the thing being described. In our description of the form of description, we make use of the analysis of Brown, in The Laws of Form (1969).

1. The Form of Descriptions

(a) Distinctions, values, motives. All things have their own form. Form may be equated with difference. If the universe were uniformly homogeneous, with no differences, nothing could be perceived;

everything would be the same. There would be no objects to be studied, no observers to examine them, no form.

Descriptions are based on distinctions. A distinction is made by an observer who perceives a difference and points it out. The difference is the basis for the distinction. Wherever a difference exists, a distinction may be made by an observer who perceives the difference. Any difference that may be perceived may be codified as a distinction.

The drawing of distinctions is an arbitrary process, governed not only by the form of the thing being described or the attributes of the observer that allow differences to be perceived, but also by the motives and values of the observer, for distinctions cannot be drawn unless the contents are seen to differ in value — and without motives, there can be no values.

Implicit in any description, then, are motives and values. Why were these particular differences singled out as important, when an infinity of other differences could be perceived instead? Because the motives of the observer defined a set of values or value differences that mattered. For example, I may distinguish between the two ends of a pencil (or between the ends and the rest) if you ask me what it is used for. "It is used for writing and erasing." But perhaps you wanted to know how much the pencil weighed, or what kind of tree the wooden part came from. These values would be relevant for other motives.

The description is related to the thing it describes in an abstract way. Whereas the form of the thing is completely manifested in the thing itself, only certain selected aspects of the form are incorporated in the description. The description is incomplete.

The form of the thing being described is continuous with the thing itself. There are no distinctions or discontinuities in the form; these are introduced by the observer, who makes distinctions. Distinctions are discrete: abstracted, codified perceptions removed from the thing being described. Since the description is based on distinctions, it too will be discrete, and not continuous, as the form is continuous. The description is imperfect.

(b) Descriptions and complex systems. Since the form of complex systems is complex, the descriptions we make of these systems may also be complex: we may base our distinctions on several kinds of differences. First, for example, we differentiate between the system and its environment. We might call the points through which we draw the distinction the "boundary." Next, we may differentiate the system as a whole into distinct parts, if these parts differ from each other and from the whole system in ways that are relevant to our motives. Furthermore, we may distinguish between different parts of these parts .... These are all vertical distinctions, where we are differentiating between units and parts of those units.

Alternatively, we may differentiate horizontally between classes or sets of parts, at any level in the system. Obviously, the level of analysis we choose will specify the units we work with, which, in turn, will determine the kinds of distinctions we make. Moreover, the description we get using one set of distinctions will not be the same as one based on a different set of distinctions. Yet both are descriptions of the same system, which has only one form. The distinctions must be equivalent in some sense.

(c) An example. Perhaps an example will help make the complexity of the situation clearer. Take a relatively simple complex system: a cupful of boiling water. Water is made of water molecules, which are made of hydrogen and oxygen atoms, and so on. The boiling water is a "liquid." It "flows." It "conforms" to the shape of its container. It is not "compressible." However, if we examine the parts of which it is made, namely the molecules, it doesn't even make sense to refer to the properties of "liquidity" or "incompressibility." You cannot "boil" or "pour" a molecule of water, although you can "boil" or "pour" a cupful of them.

What's going on here? Simply this: The property of "liquidity" with all its associated characteristics is an emergent property. It is due to the interactions of the parts of the system — in this case, the molecules. We could have described each of the molecules in the cupful of boiling water, and this description would have to be consistent with the one where we described the whole cupful instead. If we had focused on the individual molecules, we would have found nothing strange about the behavior of any single one. However, we would not have seen any evidence of "liquidity." Only when we look at the whole set of molecules, and treat the set as an individual object, would we see the object — the cupful of water — boiling.

(d) Alternative descriptions. This odd situation, where there are two equally valid — but very different — descriptions of the system, occurs whenever we examine a multi-leveled system at an interface between levels.

Because of the way the elements interact, the system one level up "behaves differently." In the example, there were emergent properties which we usually identify as the characteristics of water. What we mean by "water" in most cases is "enough water molecules to allow the characteristic emergent properties of water to be manifested." It is not that the molecules are dramatically different in the context of the more macroscopic "water"; rather, we are used to seeing large numbers of water molecules, where we can easily observe the results of their interaction on a gross scale. These results are so predominant that, in the case of water, we "see" mainly emergent properties.

This idea of alternative descriptions provides a more complete way of looking at levels. There is nothing contradictory or mysterious here; although the way in which a description at one level is related to one at another level may not always be clear. The shift we make when going from one level to another is a shift from discrete descriptions of individual elements to statistical descriptions of sets of elements. " .... [W]hile in classical mechanics the variables of system state are associated with properties of the individuals postulated by the theory, in quantum mechanics the state variable is associated with a statistical property of the postulated elements" (Nagel, 1961, p. 308). At the low level, "light" is particulate, like molecules of H_2O. At the next level up, "light" is a wave phenomenon, much like water is a liquid. These are two alternative descriptions of the same thing, each one equally valid.

(e) Transformations. How can we go from one level of description to another? The transformation from a low level to a higher one seems to require a statistical process, whereby properties of the aggregate are described. According to Nagel, we can characterize the system " ... in terms of certain statistical properties of the individual motions of the molecules. These statistical properties are represented by statistical parameters; and it turns out that a number of these parameters are associated with magnitudes of observable macroscopic properties ... " (1961, p. 291). It seems to be possible in some cases to predict what the properties of the higher level should be, given a knowledge of both the properties of the lower level elements and the way they interact.

Going up levels, the transformation seems to be determinate, in much the same way that there is always a unique mean for any set of numbers. We might call each possible arrangement of the microscopic components of the system a "microstate," and similarly, the experimentally identifiable properties of the system as a whole the "macrostates." Then we can develop a theory which " ... explains the occurrence of the macrostates of a system in terms of assumptions concerning changes in the microstates, so that the explanation depends upon the institution of correspondences between macro- and microstates. However, the correspondences are usually so specified that to a given macrostate there corresponds not one but a large number of different microstates" (Nagel, 1961, p. 314). Because of the many-to-one nature of the relation, transformations in the other direction -- from higher to lower levels -- are not so easy. Indeed, they may be impossible, in the same way that it is impossible to tell which numbers were in the set of numbers that had a mean of 7.239.

(f) Problems. The statistical nature of the relation between micro- and macrostates is not the only difficult we are faced with. The description of the microstates, like all descriptions, can only be an abstraction from the actual objects being described. Although the microstates may be continuous, their descriptions will be discrete, and thus imperfect. There may be many aspects of the microstates that go together to produce any given macrostate; some of these aspects may be left out of the microdescription. Unless the microdescription is complete, it may not be possible to derive the macrodescription from it. What is more, some of the relevant aspects of the microstate may not be evident in the case of individual elements — they themselves may require statistical descriptions.

We can begin to see how the process of description is made more difficult in complex multi-leveled systems. At each interface between levels there are alternative descriptions -- one describing the microstate and the other describing the macrostate. Usually, the description at the higher level will be simpler, more elegant. For example, different descriptions can be used for the process we commonly call "thinking." In the higher level description, we would perhaps talk about the topic of the thoughts, or the logical process that we were trying to accomplish. At a lower level, we would talk about the neurons that were firing during the process. Clearly, the description at the higher level is simpler. It is more easily done, and more easily expressed. It describes the system as a whole, rather than as a large number of parts. However, it does not provide any information about the microstates -- about the lower level elements -- the neurons.

In this case, as in most, the transformation from microdescription to macrodescription is determinate, although the one in the other direction is not. (There are many kinds of thought processes, and they all involve different sequences of neuronal firings.)

2. The Process of Observation

In order to describe a system, we need to be informed about it. The latin root for the verb "inform" -- in+form(are) -- gives a clue as to what it means to be informed about something: literally, "within" + "form" -- to have within oneself the form of the object. By "the form" of the object, we mean the basis for all the distinctions we could draw about the object, based on differences that can be perceived in the object. This is exactly what we mean by a model — a model of something has/is the form of the object it re-presents.

Unless we are merely speculating and playing with ideas, our models will be based on the results of some observations we make. In daily discourse, when dealing with common situations we are familiar with, the observation process seems to be so straightforward that it is usually ignored. If we want to know if it is raining, we look out the window. If we want to know how a car works, we look under the hood. When we want to learn how a complex system works, however, we are faced with a different kind of situation altogether. An understanding of how we make observations and what we can do with the results will help us see how we are limited in the observation situation; an understanding of the sources of limitation will allow us to develop methods that increase our observational capabilities. We begin this section, then, with an examination of the observation process.

(a) The goals of observation. When we want to understand something, we observe it. We do this because we do not know what the object does, how it works, or how it is put together. We have some uncertainty about the object. The main goal of observation is to reduce this uncertainty. When we have done so, we will have an "understanding" of the system -- we will be "informed" about the system -- we will have built a conceptual model which incorporates the form of the system. We can then use the model to predict, describe, or explain the behavior of the system.

The model is useful because it mirrors the system it represents. It is isomorphic to the system in such a way that it can be used to predict or explain what the system will do in a variety of circumstances. It can also be used heuristically, to understand properties of the system that were not evident before the model was available.

The observer will usually begin with an informal model, in which there is much vagueness and little explicit detail. This is because the observer is still uncertain as to the structure and function of the system being observed. The goal of observation is to reduce this uncertainty. When this has been done, the model will covary with its referrent -- this is what it means to be isomorphic.

(b) Observation and interaction. The reduction of uncertainty — the establishment of covariation -- is not only central to the observation process; it is the basis of the process of interaction. In order to observe a system, it is necessary to interact with it. In order to get data that indicate how the system responds in such and such a situation, it is necessary to observe it — to interact with it. How else is the uncertainty about the system to be reduced?

Aren't there situations in which we get information about a system without interacting with it? For example, I don't have to interact with the sky to see if it is raining, do I? Although the answer may seem to be stretching the point, I do. One aspect of the sky is the way it alters the light that passes through it. Unless I get in the way of some of this light, so that it lands on the sensitive parts of my eye, I cannot see the sky to tell if it is raining or not. True, this kind of observation/interaction doesn't seem to have much of an effect on the system being observed, but it is an interaction. What if I were blind? I could listen for the sound of rain. But then I am interacting with the sound waves .... Unless there is some kind of connection/interaction between the observer and the system being observed, there can be no observation, no data can be collected, and the uncertainty the observer has about the system cannot be reduced. Thus, it is necessary to interact with a system in order to observe it.

(c) Observations in complex systems. We are interested in observing complex systems. To do this, we have to interact with them to get the data. How do we proceed? Do we treat the system as a single unit, and interact directly with the whole? Do we interact instead with the parts of which it is made? Or with the parts of which the parts are made? Obviously, we begin by making some decisions. Our goals, in terms of just what it is we are trying to understand, will help us define the system in clear enough terms so we can rule out a number of approaches before we have begun. If we want to understand the structure of the system -- the ways its parts are organized -- we won't be very interested in the way the system relates to its environment, or to other systems like itself in the environment. We will, however, be very interested in factors more directly related to the organization of the system. We will look at the kind of parts it may have, for instance, and we will want to know some of the properties of these parts, to see if they provide a better understanding of the way they can be related to one another in the context of the whole system. On the other hand, if we want to see how adaptable the system is, we will want to see how it responds to environmental disturbances of different types. We will be more interested in how the system as a whole behaves in a variety of situations.

A clear statement of the goals of the investigation will help us decide both how to define the system and also how to "enter" it — how to interact with it to get the data that will be useful. The decision of how to enter the system is influenced by many factors which can best be understood from the perspective provided by the basic systems concepts we outlined earlier. Since we have a complex system, we can enter it at any of its multiple levels; that is, we can obtain our data by interacting with the system at any level we choose. Our choice of entry points will determine the kinds of data we get. This choice is limited both by the constraints that are due to properties of the system and those that are due to properties of the observer, who himself is a complex system. The next part of the section will discuss these constraints.

3. Constraints on the Available Data

In this section we will see how the goals of the observer together with certain organizational characteristics of the system are logically related to the process of observation and description in complex systems.

(a) The goals of the observer and levels of the system. When we interact with a system at a certain level, we get data which describe the system at that level. Therefore, if our goal is to understand the system at a certain level, we should enter it at that level. This is the most basic consideration when choosing an entry point. This may not be enough, however. Although data from level X may allow us to describe the behavior of the system at level X, it will usually not provide an understanding of why the things observed at level X are the way they are. To achieve a satisfying understanding of what happens at a general level, it is usually necessary to examine the system also at the two or three levels directly under the target level. Additionally, it may be very helpful to see what happens at one or two levels above the target level. This will allow the microscopic details of the lower levels to be seen in the context of the more macroscopic details of the higher ones.

Let's say the target system is a halftone photograph in a newspaper or magazine. We can understand how the different shades of grey were achieved if we examine the dots of black and white that make up the picture. But we cannot understand why the different areas in the picture were different shades of grey unless we see how they fit into the larger context of the whole picture. In that context, some shades will be highlights, while others will be shadow. All the dots together will form a single picture. Again, we won't know why that particular picture was used unless we examine it in the context of the story it illustrates. Similarly, if we want to understand the behavior of a person, we will learn how different motions are accomplished if we examine the physiology of the person; but we won't know why any particular sequence of behaviors was executed unless we put them in the larger context of the social system the person is embedded in.

(b) Emergent properties in the observation context. One reason for the requirement of looking at multiple levels is that the properties at any given level may be emergent, and thus dependent on more than just the properties of the elements at lower levels. It is much easier to work with emergent properties at the level on which they operate, than it is to work indirectly through lower levels. In order to "reach" the emergent properties indirectly, it is necessary to understand both the properties of the elements at the lower level and the way those elements are related to each other. In this way, it is possible to predict emergent properties. To do so successfully, however, is a difficult task. This is because the lower level elements may be related to each other in complex ways that may be difficult to describe or observe. Only some aspects of these relationships may influence the emergent properties at higher levels. Unless the proper subset of these aspects is included, it will not be possible to predict the emergent properties.

When the system is examined at multiple levels, however, it is possible not only to predict the characteristics of the emergent properties, but also to verify the prediction by comparing it with the actual observed properties.

So far the discussion of emergent properties has centered on the possibility of explaining or predicting them from below, with data collected at lower levels of the system. Isn't it also possible to understand an emergent property by coming at it from above, with data collected at higher levels of the system? Generally, this does not seem to be easily done. Consider what it would mean: we would be making inferences about a property at a lower level from data describing a higher one — we would be guessing blindly about the complex origins of the emergent property. Recall how these properties stem from two sources -- the properties of the lower level elements and the set of relationships between the lower level elements. Not only would we have to guess about the properties of the hypothesized elements, but also would be guessing about the way they were related to one another, even though their very existence is uncertain, since we do not have data obtained from their level.

(c) Alternative descriptions: descriptive asymmetry. We said earlier that whenever we look at a system at a hierarchical interface, there will be alternative descriptions of some phenomena — descriptions that will be equally valid, in spite of their great differences. Generally, the description at the higher level will be much more elegant and simple, but it will not provide much information about the lower level aspects of the situation. For example, I can describe an operation performed on a computer in different ways. I could say "complement the bits in word 504432. Set all the bits in the next word to zero. If the first bit in the word with the address whose value equals the contents of word 503325 is a one, complement the rest of the bits in that word ... " Alternatively, I could say "multiply the value of variable X by the value of variable Y. If the result is positive ... " Each describes a mathematical operation. Clearly, the second description is more elegant, although it gives no information about the specific bit operations that are done when the multiplication and test are done. Furthermore, although I could arrive at the second description if given the information in the first, it would be impossible for me to derive the first, given the second.

Alternative descriptions are related to each other by a statistical process. When moving from the lower level to the higher one, we shift from a focus on properties of individuals to a focus on the properties of the aggregated individuals. The description of the aggregate is the higher level description. Although it is possible to transform low level data to higher level data, it is not possible to go in the other direction -- given the characteristics of an aggreagate, it is generally not possible to derive those of the individuals in the aggregate. For this reason, data from levels lower than the target level are more valuable than data from higher levels.

(d) Summary. To summarize, it is desirable to obtain data by interacting with each of several levels in the system being observed — both from the levels immediately below the target level and from the levels directly above the target level. Because of the way the low levels are related to higher levels, in terms of both functional arrangement and descriptive characteristics, data from lower levels is often more useful and important than data from higher levels, although the investigator's task is made easier if data are also available from higher levels. If data are only available from higher levels, it may be impossible to make any definitive statements about lower-level phenomena.

4. Dealing with Constraints -- Tools

The main difficulty in the study of complex systems, especially when communication or "information processing" are involved, is that data are often not available from the appropriate levels. It was an understanding of the logical nature of the hierarchical interface that made this difficulty clear, and it is through this same understanding that we can develop methods that will allow us to go past some of the limitations we are currently faced with. Once this has been done, we can begin to solve some of the problems that more traditional methods fail to recognize.

We find ourselves faced with two kinds of situations. In the first, we have data from a level higher than the level we wish to understand. The best we can do in these cases is guess about the lower-level phenomena in which we are interested. The reason for this has been discussed earlier. In the second case, we have data from a lower level than the one we wish to understand. Many situations involving complex systems, especially information processing systems, are of this type. We consider two characteristics of the hierarchical interface: (a) First is the statistical nature of the relationship between lower level descriptions and higher level descriptions of the same phenomena, which suggests that any analytic methods designed to provide high level descriptions will use statistical methods, rather than precisely exact mathematical solution-finding approaches. The exact nature of the statistical transformations that should be used, however, is not clear. In some cases, pattern-recognition techniques seem to work best. At any rate, more theoretical work is needed if we are to understand fully the logic of alternative descriptions.
(b) Second is the fact that emergent properties must be examined in the light of the relationships between the lower-level components of the system, which suggests that any methods designed to explain emergent properties at higher levels will examine the relationships between low-level units. Again, we do not yet know how to tell which properties of the low-level parts or which aspects of the relationships between the low-level parts are important in determining the emergent properties. This area also needs more theoretical exploration.

At any rate, we know that analytic methods will have to look for statistical solutions to alternate descriptions and includde the network of relationships among the parts of the system for explanations of emergent properties, such as "information," "communication," or "control" processes.

CHAPTER FIVE

STRUCTURE: THE APPEARANCE OF FORM

Chapter Three was entitled "The Form of Systems." Chapter Four was about "Access to Form: The Processes of Observation and Description." In those chapters we were asking the questions, "What are we looking at?" and "How do we look?" In this chapter we ask "What will form look like?" The "form" in this question is the same "form" we have been discussing all along. Its appearance is structure; structure is how form which has been observed is described. We will examine the meaning of this statement in some detail in later sections of this chapter. The discussion there builds on the ideas of the previous two chapters, so a review of the main points discussed there will be useful in setting the stage for our discussion of structure.

REVIEW

In Chapter Three we talked about some general characteristics of complex systems. Although most of the discussion there was concerned with the development of a clear technical vocabulary through the explication of a set of critical concepts, we also touched on some basic properties that may be observed in complex systems. It is these properties that are of interest here. If any one statement about systems in general is fundamental, it is this: The properties of a system are due both to the properties of its parts and to the way the parts are related to one another. In order to understand the properties of a system, then, it is necessary to see not only what kind of parts it is made of, but also how those parts interact with one another.

A process fundamental to the functioning of systems is the interaction process, which we discussed in terms of constraint. There is constraint whenever there is interaction or influence. Constraint is the flip side of independence. If a set of objects are independent of one another, they do not interact with each other. The degrees of freedom of the set is equal to the sum of the degrees of freedom of the elements in the set. However, if the elements do interact with one another, they will not be independent. The degrees of freedom in this kind of set will be less than if the elements were independent. This reduction in degrees of freedom is constraint.

We also saw in Chapter Three that if a system is to have properties different from the properties of its parts -- we called them emergent properties -- the parts may not be independent of one another. A set of independent parts shows no emergent properties. Since the amount of interdependence is directly (in fact, identically) related to the amount of constraint (i.e., to the size of the reduction in degrees of freedom), the amount of constraint can be used as an index of the likelihood that the system as a whole will show emergent properties. A system with no constraint is a system of independent units. Such a system shows no emergent properties. It would be better to study the individual parts making up such a "system" than to study the system as a whole.

Thirdly, we saw in Chapter Three how information is related to the systems which must process it. There can only "be" information in sufficiently complex systems, for the "meaning" of information is "there" only because of the way the system processes it, and we viewed information processes as emergent processes, requiring multiple levels in the system. The arbitrary aspect of information coding -- the "meaning" -- is not a function of the signal which is transmitted in information exchanges; rather, it is a function of the changes in the system as the signal is transformed up or down levels in the information process. The exact nature of these changes can be explained only if the set of relationships between the parts of the system at each level in the system is known. The difficulty of gaining this kind of information is one of the points brought out in Chapter Four.

In Chapter Four we talked about the processes of observation and description. Chapter Three covered the form of systems -- how they are put together and how they work -- and Chapter Four covered the way in which we may gain access to that form. We began by describing descriptions. We said that all descriptions are based on arbitrary distinctions, drawn by the person describing the system. In order to draw distinctions we need some way of getting information about the system. This is accomplished through the process of observation, where the observer interacts with the system in such a way as to reduce the uncertainty he has regarding some aspect of the system. All observations require interaction with the system being observed.

Now, in Chapter Three we saw some of the limitations on interactions between two systems. An interaction involves the establishment of covariance between the two systems. This covariation may take place at any level common to the two systems. The ability of the observer to establish and interpret this covariation in such a way as to answer the particular question that was asked will depend in part on the choice of levels at which the system is "entered." This is a complex issue, because of the relation between descriptions at different levels and the behavior of the system at different levels.

Here the logic of alternative descriptions becomes crucial. Very simply, the logic of alternative descriptions says that there may be many descriptions of the same system, each of which is framed from a different level of analysis. We may speak about the whole system, or its parts, or their parts, and so on. The complicating issue here is the emergence of apparently new properties as we move up to higher and higher levels of analysis. Thus, although a macrodescription must be equivalent to a microdescription of the same system, this equivalence might not be apparent. Only if the microdescription is complete in its inclusion of all the relevant aspects of the parts of the system will an equivalence be possible. The equivalence will only be seen if the microdescription is properly transformed, perhaps by statistical or pattern-recognition techniques, into an alternate macrodescription. The choice of aspects to be included or excluded is of primary importance, as is the actual transformation method that is adopted.

If a transformation is formulated and applied to a microdescription, the results may be compared to a valid macrodescription as a test of the fitness of the transformation. If a match is obtained, the transformation may be used as a model of the system, in terms of part-whole relationships. This is a way of checking to see if the important and relevant aspects are being included in an explanation of how the system works.

The process of moving from micro- to macrodescriptions was covered in some depth because it was felt to be important for two reasons: First, it provides a better conceptual understanding for the relationship between a description of a system and a description of the parts of which the system is made. The relationship here is one of structure to function, where structure is a macro-concept with a parallel micro-concept of function. This comparison will be used again in later chapters. Second, the process provides a way of working with systems where data are only available from low levels, while descriptions are desired at high levels. It suggests the basic type of transformation that should be used to make the most out of the data. Specifically, the transformation must be statistical, where there is a loss of a great deal of microscopic information (which is useless at a macroscopic level anyway) in exchange for a gain in macroscopic information. It is this selective loss of detail that allows macrodescriptions to be useful. The principle of selective loss of detail is not only important to observers of systems; it is crucial also for the operation of the system itself, as decision-making units strive to control the behavior of the system and to monitor their effectiveness by means of feedback mechanisms of various types. The parallel importance of the concept both to the system and to its observer suggests that the approach to the observation process which we have outlined here is a coherent one -- the observer is bound by the same rules as the system being observed.

PREVIEW

As the title to this chapter implies, we are examining structure as the appearance of form. Structure is what the observer sees, while form is what the system is. We can never have direct access to form — all knowledge of form is indirect, achieved through the process of observation. Thus, structure -- the "image" of the form built from information obtained through observation -- is not the same as the form; it is the appearance of form. In this chapter we explore the concept of structure as it applies to the general case of hierarchical systems. In this explication we adopt a strategy that starts with these concepts:

(1) A basic definition of structure as deviation from chaos or randomness;

(2) A systems rationale which says that emergent properties of the system (including structure) are due to both the properties of the parts of which the system is made and to the way those parts are related to one another; and

(3) The concept of alternative descriptions, which suggests how to go from a description at a low level to one of a higher level.

We combine these basic concepts to obtain a notion of system structure which is solidly based in the theory of observation and description in complex systems. The practical result of this logical process is a set of guidelines that suggests both what to look for at higher levels of analysis and how to proceed, given only low-level data.

We begin with a discussion of the definition of structure as deviation from chaos or randomness. After the basic notion has been clarified, we move on to an analysis of structure from the perspective of complex, multi-leveled systems. In this section we examine the relation between structure and form, tying in the ideas central to the processes of observation and description. Constraint plays a central role in the logical analysis, and in the next section we suggest that it also is central in the analysis of complex systems. In this Part we outline some procedures that may be used to approach any complex system when the structure and organization are not known. A flowchart for a general "analytic algorithm" is presented and it is argued that this approach is appropriate for the general case.

STRUCTURE

We will define structure as a deviation from randomness or chaos. In order for the definition to be useful, we need to know precisely what this "randomness" is, since it is the major term in the definition. Specifically, we will be interested in the nature of random, or unstructured, systems. Thus, we begin with randomness and random systems.

Structure and Randomness

In a random or unstructured system there is no pattern. The parts of the system are independent of one another. All possible states of the system (and its parts) are equiprobable. All of these characteristics are important. Important also are some other characteristics implied by randomness. Since the parts of the system are independent, there is no systematic covariation of the parts. When one part is known to be in the system, the discovery that other parts are also in the system has no effect on our expectations of the behavior of the first part. The other parts are not related to one another: they are not constrained as a result of being in the system. For all practical purposes, there is no discernable difference between a part in isolation and one in a random system: the parts in the random system are free to behave in the same ways they would behave if they were not in the system. The total variety of the random system is exactly the sum of the varieties of the parts. Furthermore, a random system is best described as a collection of independent parts at the level of the parts.

Randomness is thus the absence of pattern, of order, of interdependence, of covariation, of constraint, of predictability. A random system is not in principle different from a set of randomly selected parts. In fact, many theorists would go as far as to say that there are no "random systems." They would say that what we have been calling "random systems" are not really systems at all. According to this viewpoint, the concept of a "random system" is self-contradictory, since the characteristic which distinguishes a system from a mere aggregation is the interdependence of the parts of the system. In an aggregate, the parts are independent. This is a definitional problem. Here we are using the looser definition of system, which allows the existence of "random systems."

A structured system is one that is not random. In a structured system there is pattern, regularity, order, organization. The parts are, at least to an extent, interdependent. Since they are interdependent, there is some covariation of the parts. Thus, a knowledge of the states of some parts is likely to tell us something about the states of other parts. The parts are thus constrained by being in the system -- inside the system some situations are more likely to occur than outside the system, while other situations which are seen outside the system are not seen inside the system.

In a structured system the relationships between the parts are not random. Thus, the existence of some relationships is more certain than the existence of others. A knowledge of some relationships will tell us something about the probability of the existence of other relationships -- some will be more probable, while others will be less probable.

Structured systems are thus different from random or unstructured ones. Let us examine the deeper relation between structure and some of the other concepts we have been working with.

Emergent Properties, Levels, and Structure

In a random or unstructured system, the parts are independent of one another. What relationships there are between parts are also random -- the relationships are independent of one another. There are thus two "levels" of randomness. At the first, the parts are independent of one another. There is no interaction among any of the parts.

At the second level, there may be interactions among the parts, but these relationships are independent of one another. Many times there will be a mixture of the two types of randomness. For example, in a social system that includes, say, the people in a waiting room, most of the people will not interact with any of the others. What interactions there are are essentially random -- they are not coordinated in any special way. Throughout the rest of the discussion, when we say "random system," we will mean random at the second level.

If relationships are viewed as constraints, this means that the constraints in a random system are random also. There is no relation between any one constraint and any of the others.

We said earlier that emergent properties have their roots in the relationships among the parts. "Because of the relationships among the parts," we said, "the system behaves in ways the individual parts do not or could not." We also said that the notion of emergent properties is very closely tied to the notion of levels, which we defined in terms of emergent properties: a set of interacting parts is said to be at a higher level than the individual parts if the set displays emergent properties as a result of the interaction.

Two notions stem from this discussion. First, in a random or unstructured system there are no emergent properties. This, of course, is because of the independence of the parts and the absence of relationships between them. Second, in a random or unstructured system, there is no new level above the level of the parts. In other words, the level of the system is the same as the level of the parts. The properties of random or unstructured systems are thus the same as the properties of the parts of those systems, which behave simply as their parts behave. In other words, the behavior of random systems is the same as the behavior of their parts -- no more and no less.

Structure and Form

We said earlier in this chapter that "structure is the appearance of form"; that "structure is what the observer sees" and "form is what the system is." This section expands on these points.

The form of the system is what the system is. The form precedes all the distinctions that could be drawn, because it incorporates (is) all the differences that can be used as bases for distinctions. The form defines the system. The form is the system ...

If the system is formless, no distinctions can be drawn because there are no differences. Nothing can be said about a formless system, other than that it is formless. The more differences there are in the form, the more distinctions can be drawn, and the more that can be said about the system.

Now, the observer does not have direct access to form. The best the observer can do is "sketch in" the form, by interacting with the system in such a way as to discover some of the differences in the form. In this way, it is discovered that systems have boundaries, or are hollow, or are red, and so on.

If the observer is uncertain as to the aspects of form that are likely to be seen, he is likely to make a great number of observations in the hopes that it will be possible to discover the important differences in the form by looking indirectly -- at the data points provided by the observations. If the data points are all the same, no differences are observed and very little can be said. There is no variety in the form.

On the other hand, if the data points are all different, the only description that is faithful is one that reports all the differences. Here the best description of the system is a description of its parts. This situation is the random system -- chaos. There is no constraint.

These two types of systems -- organized simplicity (no variety) and chaotic complexity (no constraint) were discussed earlier. We are not interested in these types of systems, but rather in the third class -- organized complexity -- where there is variety and constraint.

In this type of system there will be some systematic variation or pattern instead of absolute chaos in terms of differences in the form (and hopefully the observations), so that some sets of parts are different from other sets of parts, or some sets of relationships are different in some way from other sets of relationships. It is possible to base distinctions on these higher-level differences in the system. Here the best description of the system is not one which merely distinguishes each part from all the others. Rather, it will be one which distinguishes different parts, but not similar ones. Thus, the description of the system will be simpler than the combined descriptions of all the parts in the system.

What is being described here is the structure of the system. The term "structure" refers to the appearance of form -- the differences reported by the observer in the data points. In a sense, structure is the form as reconstructed or modelled by the observer.

Properties, Descriptions, and Constraints

The concepts of properties, descriptions, and constraint are all intimately related. "Properties" are seen in descriptions of constrained systems. How does this work?

Descriptions, we saw earlier, involve the drawing of distinctions. In order for a distinction to be drawn, an observable difference must be perceived by an observer. Properties are simply a type of distinction drawn on the basis of some differences in the form of the system being described. For example, if we speak of a system's weight or age or length, we are distinguishing between the particular system observed and all systems having different weights, lengths, or ages. These distinctions involve the establishment of relations between the system being observed and other systems. If we speak of the shape of a system, or talk about the relation of one side of the system to the other side, we are distinguishing between different parts of the system. Here we are concerned with the establishment of relations between the parts of the system. These relationships, like all relationships, depends on some kind of constraint. If there were no constraints, differences could not be observed, and distinctions could not be drawn. The system would have no observable properties.

The more differences we can observe, the more we can infer about constraints operating in the system, and the more properties we can describe of the system. The more random a system is, the less constrained it will be, and the less differences its form will have. This means that we can make fewer distinctions, and describe fewer properties. We would say that the system "has" few observable properties.

A system may be composed of a set of unconstrained or randomly organized parts. Even if these parts are themselves highly structured (i.e., if they show many clear properties), such a random system will "show" no properties not seen in the parts of which it is made. The best that can be done to describe the "properties" of the system in this case is to describe the properties of its parts. The only constraint operating in the system is in the individual parts.

If there is some constraint in the way the parts are organized, however, this will lead to differences which can be observed and described as "properties." Snakes, for example, have two ends. The ends are different from the middle, and one end is different from the other. These differences are due to constraints in the way the cells of the snake are organized. If these constraints were not allowed to operate, the ends would be the same and it would not make sense to speak of the "head" or the "tail." If the ends were not different from the middle, the snake would be like a hoop, with no ends of any kind. It is the constraint that allows differences which determine the properties of the system.

Emergent properties, like other properties, are based on differences/constraints. When the parts of a system join one another in the relationships implied by membership in the system, they are constrained in some way. It is this additional constraint that allows emergent properties to "be." The more organized are the relationships among the parts, i.e., the more related/constrained are the relationships throughout the system, the more the system as a whole will show properties as a unit. On the other hand, if the relationships among the parts are unconstrained -- independent of one another -- there may be no differences to be observed and thus no emergent properties. The system "behaves" as a collection of randomly organized parts.

Even if the parts are constrained by their relationships with other parts, there may be no constraint in the overall set of relationships. This situation is illustrated by a bucket of sand. While each grain is physically constrained by its neighbors, there is no overall order to the set of these physical constraints. Because there are no constraints on the set of constraints, there are no differences over the set and, therefore, the sand as a whole shows no emergent properties.

Sources of Constraint: Structure vs. Control Systems

There seem to be two sources of constraint on the elements in complex systems. The first is simple -- the physical characteristics of the elements directly limit their interactive capabilities. To understand how an element is bound by this type of constraint it is necessary only to examine the element itself. This is because the constraint is due to the form of the element, and not to other considerations.

The second source of constraint on the elements is more complex. Because of the way the parts interact, the system may have some emergent characteristics that modify the behavior of the parts in the system. This amounts to what Pattee (1973) called a "feedback loop between levels." Pattee argues that this kind of constraint is what allows the system to control its own behavior. In order to understand how the system constrains the elements of which it is made, it is necessary to examine the whole system. This is because the constraint is due to the form of the system, rather than the form of the parts.

In the systems where only the first kind of constraint is present, there is a hierarchy of organization, where the system is made of parts which are made of smaller parts, and so on. These hierarchies are also characterized by hierarchies of numbers, forces, and time scales, so that larger numbers, weaker forces, and longer time scales are associated with higher levels in the system (Pattee, 1973, pp. 75-77). These systems can be described by dynamic equations written for one level of the system at a time. In these equations it is possible to make the simplifying approximation that there is a "typical particle" which is characteristic or representative of the collection at that level. The behavior of gaseous systems is so described with the model of perfectly elastic molecules, where the individual details are averaged out and dynamics of higher levels are constant.

When the upper levels exert partial constraints on the details of behavior of lower levels, however, the case is entirely different. In this situation, there will be a hierarchy of controls, instead of a simple hierarchy of structures. Pattee says that:

> In a control hierarchy the upper level exerts a specific, dynamic constraint on the details of the motion at lower level, so that the fast dynamics of the lower level cannot simply be averaged out. The collection of subunits that forms the upper level in a structural hierarchy now also acts as a constraint on the motions of selected individual subunits. This amounts to a feedback path between levels. Therefore, the physical behavior of a control hierarchy must take into account at least two levels at a time, and what is worse, the one-particle approximation fails because the constrained subunits are atypical. (Pattee, 1973, p. 77)

The constraint on the lower levels will be exerted by the higher levels, rather than only by other lower-level phenomena. One effect of this kind of constraint will be that the relationships among the lower-level parts will be organized into higher-level patterns. The relationships themselves will be organized in ways that would only be possible with direction from higher levels. In order to see how the parts behave it will be necessary to examine them all in the context of the overall system because there are no "typical particles" in this kind of system. This analysis requires an examination of two levels at a time.

Summary

Before going into a description of general research approaches to complex systems, it will be useful to summarize the important points concerning constraint which have been brought out in the discussion of this chapter.

(1) Structuring in systems results from constrained variety. Thus, random systems are unconstrained.

(2) Properties may be described when differences resulting from constraints are observed.

(3) Emergent properties are dependent on organizational constraints, which operate on the level of relationships among parts.

(4) Total constraint results in static structures or organized simplicity. Zero constraint results in chaotic aggregates. Neither of these allow for interesting emergent properties.

(5) Constraint exerted on low levels by low levels allows the development of local structures but not system level dynamics.

(6) Constraint exerted on low levels by higher levels allows the development of complex forms of organization where hierarchical control processes are seen. This kind of constraint will be evidenced by the organization of relationships between parts into higher-level patterns which can only be examined in the context of the entire set of relationships.

STRATEGIES FOR RESEARCH

The goal of much research in the area of complex systems is to gain an understanding of the characteristic properties of the system together with an explanation of how they come to be manifested by the system. In the process of achieving such an understanding, a number of questions will have to be answered. One of the first of these will be: Are the observed properties associated with the system as a whole or with the individual parts of which the system is made? If the answer is not immediately obvious, due to the fact that the particular properties observed could only exist in a system of several parts, the question can be answered empirically, by studying individually the parts after they have been isolated from the rest of the system. If the properties are no longer observed, they are associated with the system.

In many cases, however, it may not be clear what the "properties" of the system are. This could be due to the fact that the particular system being observed was never examined as a whole intact system before. It may be due to the fact that techniques for measurement or observation have not been developed for the particular type of system being studied. It may be due to the fact that virtually nothing is known about the system at all, other than that it contains some particular set of member elements.

If the last is the case, it will be necessary to approach the system from a most basic perspective. This approach might be one that says to get the best possible information about the system and use as much of this information as can be utilized in formulating a description of the system.

We said in Chapter Four that, while it is possible to make high-level statements about systems when only low-level data are available, it will probably not be possible to move in the other direction. Because of this limitation, we will deal here only with cases where low-level data from sets of elements which we wish to study as systems are available.

The kind of information that will be most useful is information descriptive of the parts of the system and of the relationships among those parts. Given this kind of data in a situation where we don't know anything about the system as a whole, the first question we would ask is, "How constrained is the system?" Here we would be asking how interdependent are the parts and their relationships. If the system is a random one, there is no need to study it as a system. It will be better to study it as a collection of separate parts. If the system is constrained, however, we will want to know more about the way it is organized. How many levels are there? What kinds of constraint are operating on the lower-level parts? What emergent properties can be identified? What are the dynamics of the system? What kinds of control mechanisms does it contain?

In the next section we present a flow chart for a "general algorithm to study structuring in complex systems" (GASSICS). The process described there will result in a description of the system, in terms of hierarchical levels, with a specification of the organization of parts at each level. This is the first step toward gaining an understanding of the system. Although working through the process will give a description of how the system is organized hierarchically, it will not explain the dynamic behavior of the system. In order to do this, it will be necessary to study the system as it interacts with its environment across many periods of time.

A GENERAL ALGORITHM FOR STUDYING STRUCTURE IN COMPLEX SYSTEMS

We present here a general research strategy for studying structuring in complex systems. Because the strategy can be outlined into an explicit set of steps, we call it an "algorithm." Compared to many "approaches" or "methods," this one is almost mechanical to carry out. We call the procedure GASSICS -- General Algorithm for Studying Structure in Complex Systems.

The GASSICS procedure was formulated to take advantage of the conceptualization we have presented in this Part. All the basic ideas about form, constraint, structure, the processes of observation and description, and the logic of alternative descriptions are incorporated into the procedure.

In the last section we outlined some general research questions we would ask when approaching an unknown complex system. The first questions centered on the form of the system, or rather the appearance of the form to an observer — structure. Since we viewed structure in terms of constraint, constraint is a core concept in the GASSICS procedure. The final result of a complete application of the procedure is a description of the system in terms of hierarchical levels, with a specification of the organization of parts at each level.

The procedure begins with a low-level description of the parts of the system. In general terms, these parts would be at the highest level at which units are both readily identifiable and accessible for observation. The procedure itself is organized as a recursive loop. Each cycle through the loop moves up one level of analysis, so that the result of one cycle is a description of units one level higher than the original parts. These units are used as parts in the next cycle through the loop. There are two points in each loop where the procedure may terminate. One of these will happen if, at some level, the system is found to be random or unconstrained. The other will happen where the highest level -- that of the whole system -- is reached. Let us examine the procedure more closely.

The Loop

We begin with a low-level description of the parts of the system. This description must include data about the relationships between the parts, in terms that can be related to constraint. When two parts are related, one or both is constrained by the relationship. The data, then, will be a description of each relationship between parts. (The data are described more fully in the next chapter.) The collection of this data is the first step in the loop.

The second step in the loop is to see if the set of relationships among the parts is constrained or non-random. If it is, there is evidence of structuring, in terms of higher-level organization. This implies either that the system as a whole may be broken down into differentiated parts or that the next level above the level of the parts is the level of the whole system.

The third step, of course, is to identify the differentiated parts, if there are any.

When the parts have been found, we begin another loop by obtaining information about relationships between the new, second-level parts. The relationships at this level are not likely to "look" the same as the original low-level relationships. This would be expected because the units we are dealing with are not the same -- in fact, they are made up of several smaller units. This problem is discussed in more detail in Chapter Six.

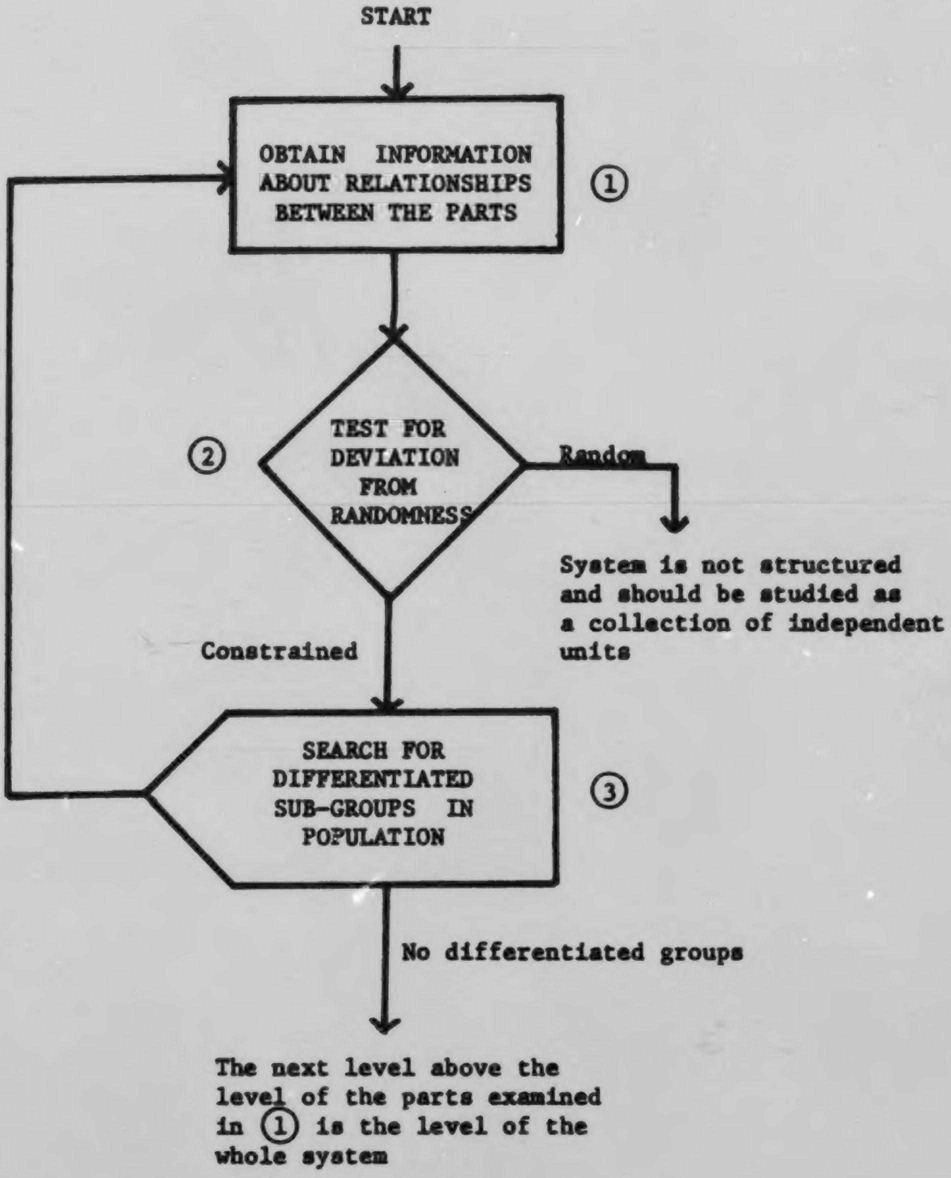
At any rate, this new set of relationships would be examined for constraint, just as the earlier set was. The rest of the loop is completed in an analogous manner each time, the only difference being the level of analysis each time through the cycle.

The GASSICS procedure is diagrammed in flowchart form in Figure 1. The three main steps in the procedure are shown by the numbered parts of the diagram. The box at the top is the data collection step; the diamond is the constraint or structure test; and the trapezoid is the step in which differentiated parts are identified.

The first three chapters of the next Part correspond to the three steps in the GASSICS loop; the fourth presents a general discussion of the procedure.

Figure 1

GASSICS: A GENERAL ALGORITHM FOR STUDYING STRUCTURING IN COMPLEX SYSTEMS


START
OBTAIN INFORMATION ABOUT RELATIONSHIPS BETWEEN THE PARTS
①
②
TEST FOR DEVIATION FROM RANDOMNESS
Random
System is not structured and should be studied as a collection of independent units
Constrained
SEARCH FOR DIFFERENTIATED SUB-GROUPS IN POPULATION
③
No differentiated groups
The next level above the level of the parts examined in ① is the level of the whole system

PART THREE

OPERATIONALIZATIONS

In the chapters of Part Two we discussed the conceptual basis for a coherent approach to the study of complex systems. We started with an analysis of the form of systems and moved on to an examination of the processes of observation and description. From there we went on to discuss structuring in complex systems — what would we find, given an understanding of what we were looking at and how the "looking" process works. Finally, we outlined GASSICS, a general procedure to use when approaching unknown systems.

The chapters of Part Three present an operationalization of one cycle of the GASSICS procedure for use in large-scale systems.

The steps in a cycle of the GASSICS loop are:

(1) Obtain information on the relationships between the parts of the system.

(2) Examine those data to determine whether or not the parts are organized in some structured manner. If the set of relationships between the parts is found to be structured or constrained ...

(3) ... examine the set of parts for units one level higher than the original parts.

If the total set of parts breaks down into several higher level units, go back to the first step, using the units as parts.

Chapter Six describes measurement problems in general and gives specific examples of a technique designed to collect the appropriate data for the study of large social systems. In Chapter Seven we develop the statistical tools needed to perform the test of structure used in the second step of the GASSICS loop. Chapter Eight presents an algorithm that identifies units one level above the level of the parts, together with a description of a computerized implementation of the algorithm. Chapter Nine consolidates the earlier chapters of this Part and discusses this operationalization of the GASSICS loop in general terms. In this overview, we see how the analytic procedure fits into the overall approach to complex systems, as well as how it relates to other methods of analysis.

CHAPTER SIX

STEP ONE OF GASSICS: DATA COLLECTION

INTRODUCTION

In this chapter we will discuss the first step of the GASSICS procedure. As we saw at the end of Chapter Five, the procedure is a loop, with one cycle for each level in the system above the level of the parts. We begin with information about the relationships between the parts and move up levels, one at a time, until we come to the highest level in the system.

Assumptions

There are some assumptions being made which should be discussed at this point. First, we only move up levels. The reason we go up instead of down is, of course, the asymmetrical relation between alternate descriptions at different levels. This issue was discussed in Chapter Four.

The second assumption we make is that levels are discrete, and can be taken one at a time. If this were not the case, we would not be able to have a discrete, repetitive loop as we do in the GASSICS procedure. We would not be able to have any single description without having every description -- there would only be one description -- with none of the "alternate but equivalent" business we spent so much time on in Chapter Four.

How do we justify the second assumption? There are two lines of thought we can follow. The first is based on the nature of descriptions and the second is more empirical. Let us start with the logical argument first. The argument will be that the very nature of descriptions demands a discrete organization into levels. It goes like this: descriptions are based on distinctions, which are made when differences are noted. If there is a difference perceived in the form, one part can be distinguished from the other. Distinctions cannot be made when differences cannot be reliably noted. Thus, distinctions are discrete.

Since descriptions are based on distinctions, they too must be discrete.

What is the difference, in terms of distinctions, between two alternate descriptions at different levels? Basically this -- different distinctions are being made. Recall that we defined the concept of levels as a concept of description. We said a set of interacting parts would be at a higher level if we observed properties in the interacting set that we could not observe in the set of parts taken one at a time. We called these properties emergent properties and said they were due to the interaction of the parts.

At higher levels, then, we will be describing emergent properties or characteristics of the system, rather than simply properties of the parts. We discussed properties in terms of distinctions based on differences in the form. These differences were equated with constraints in the system. At higher levels, there are new properties -- new distinctions, based on new differences. These differences result from constraints due to the way the parts interact.

This is all fundamentally discrete. A description at one level will take account of one set of distinctions, while a description at a higher level will take account of a different set of distinctions, based on the only thing that is different between a collection of isolated parts and a system made of interacting parts -- the interactions.

Thus, we have discrete levels of description. As we include more and more global interactions -- that is, interactions covering or constraining larger and larger segments of the system -- we move up to higher and higher levels.

The empirical argument for discrete levels is based on the idea of near decomposability, a concept described by Simon (1973, pp. 9-10). If we examine the kinds of interactions binding the parts of a physical system, we see that there are sharp changes in the strengths of these bonds as we move up or down levels. For example, protons and neutrons interact primarily through what is called the "strong force." At this level, bonds are on the order of 140 million electron volts each. One level higher we have molecules, where the electro-magnetic forces are on the order of five to six electron volts. At this level the strong forces do not operate. The bonds responsible for the tertiary structure of large macromolecules are on the order of one-half of an electron volt. "It is precisely this sharp gradation in bond strengths at successive levels that causes the system to appear hierarchic and to behave so" (Pattee, 1973, p. 9).

Not only are there different kinds of bonds at different levels, but there are also different behavioral characteristics. Perhaps most fundamental here is what Simon calls the "associated frequencies" of each level.

> Motions of the system determined by the high frequency modes will control ... the internal interactions of the components of the lower level subsystems in the hierarchy, but will not be involved in the interactions among those subsystems. Moreover, these motions will be so rapid that the corresponding subsystems will appear always to be in equilibrium and most of their internal degrees of freedom will vanish. In their relations with each other, the several subsystems will behave like rigid bodies, so to speak.
>
> The middle band of frequencies, which remains after we have eliminated the very high and very low frequencies, will determine the observable dynamics of the system under study — the dynamics of interaction of the major subsystems. (Simon, 1973, p. 10)

Because there are such sharp breaks as we go up or down levels, Simon calls these systems "nearly decomposable," meaning that for almost all practical purposes the system can be "decomposed" into discrete levels. Thus, the second assumption is supported by the empirical evidence, as well as logical arguments.

Simon's presentation of near decomposability was discussed in terms of physical systems — electrons, atoms, and so on. The same general concept seems to apply to social systems as well. For example, the relation between a husband and wife is usually much stronger than relationships between different families in a neighborhood. It is also a fundamentally different kind of relationship. Going up another level, the relationships between families in a community are different from the relationships between communities, cities, states, or countries. The same kind of hierarchy can be seen in large organizations, as we move from work teams to departments to divisions, and so on. The implications of the second assumption are discussed in the next section.

Implications of the Assumption of Discrete Levels

The assumption that levels are discrete suggests that we can go up one level at a time. The reasons behind the discrete nature suggest the kinds of differences we should expect to see as we go up levels. Both the logical and empirical arguments discussed above imply that the relations between the parts at one level will be different from the relations between the parts at different levels. One aspect, however, will be the same -- all relationships imply some kind of covariation between the parts. All relationships constrain one or both of the parts, no matter what level the parts are at.

RELATIONSHIPS

What more can we say about relationships at this point? First, relationships will "look different" at different levels. They will require different observational techniques to be used. Although it is possible that a relationship between units at one level will be equivalent to (or reducible to) a set of relationships between the parts of those units at a lower level, there is no reason to expect this to be true in all cases, or even in a majority of cases.

Second, the exact nature of the relationships that may be entered by a unit at some level will be mostly a function of the unit at that level. In other words, it will be difficult to determine the nature of relationships at some level without some information about how those units behave, even if information about the parts of those units at lower levels is available. The lower level information may be very useful, but it might not be sufficient in all cases. It will depend on the system.

Third, there are some general characteristics that can be used to describe relationships at all levels. These general characteristics are described in the next section.

General Characteristics of Relationships

1. Strength

Probably the most basic aspect of the relationship is its strength. How much constraint is there because of the relationship? To what extent are the behaviors of the involved elements influenced?

2. Symmetry

Are one or both of the involved elements influenced? If both are influenced equally, the relationship would be symmetrical. In symmetrical relationships, there is no concept of direction -- the elements are mutually influenced. If only one element is influenced, the relationship would be asymmetrical, and it would make sense to speak of the direction of the relationship, in terms of which element influences the other.

3. Transitivity

Does the influence of one relationship carry over to the other relationships the elements may be involved in? In other words, if A is related to B and B is related to C, does it follow that A must covary with (be related to) C? If this is the case, the relationship would be transitive. Otherwise, it would be intransitive.

Matter-energy vs. Information Relationships

In Chapter Three we discussed the differences between matter-energy and information relationships. One aspect of the difference we did not discuss was this: in systems where the interactions are all matter-energy, there will usually be only one type of relationship between the parts at any one level. For example, in physical systems, at the level of subatomic interaction between the protons and neutrons of an atom, only the strong force has any effect. The electromagnetic forces, weak forces, and gravitational forces have no effect at this level. At the next level up, where we are speaking of bonds between atoms in molecular structures, the only forces that are important are electromagnetic. Thus, at any level only one kind of interaction may be observed between parts.

In information systems, however, this may not be the case. It is possible for there to be many different types of relationships concurrently between the parts at some level. In fact, it is possible for there to be several distinctly different systems, all composed of one set of parts. In this "overlapping" system situation, the parts at some levels may have several "modes" of behavior, so that if these modes are examined separately, each will appear as a distinct, complete system. It seems likely that these multiple modes are more accurately described as different aspects of the microscopic characteristics of the low-level parts. Some of these aspects may be important at the low levels of the parts, but irrelevant to the system as a whole. The resolution of this issue may be primarily an empirical task, but it surely demands more in the way of logical clarification, too. At the present time, it seems advisable to isolate these modes during analysis, at least when using this first operationalization of the GASSICS procedure.

The discussion so far in this chapter has been conceptual. In the next sections we turn first to some issues related to operationalization in general, and second, to a specific application.

OPERATIONALIZATION IN GENERAL

The issues we have been discussing are conceptual. In order to use them we have to translate them into operational procedures. In other words, we need to build a model of the system, in the form of data. The data will be the operational counterparts of the conceptual terms. In order to preserve the conceptual clarity we have been trying to maintain, we observe the distinction between the conceptual and operational components. To underline this distinction, we introduce new terms to refer to the operational aspects. Sometimes the operationalizing procedures will lead to situations where new, additional distinctions need to be made, due to the relation between the conceptual and operational systems. Both this new terminology and the distinctions associated with it are presented in this section.

We start with the basics. Where we had the concept of the system, we use the operational network. Where the system was made of parts and relationships between them, the network is a set of nodes with links between them.

The network is not the system. It is the image or model we use of the system, in the form of data. Similarly, the nodes are not the parts of the system; rather, they are artificial constructs that represent or stand for the real parts. Finally, links are not relationships. Instead, a link between a pair of nodes indicates that there is a relationship between the corresponding pair of parts of the system.

Symmetry/Reciprocity

Just as we use nodes and links instead of parts and relationships, we also differentiate between the theoretical concept of symmetry and its operational counterpart, reciprocity. In this case the distinction is most important when measurement techniques may be unrealiable. For example, if a relationship is conceptualized as symmetrical, if A is related to B, B must also be related to A. This is analogous to the case when both parts are equally constrained by the relationship. Here we would expect a parallel symmetry in the data.

If, on the other hand, the constraint is not equal, so that one node is constrained while the other is not, we might find that A is related to B but B is not related to A. Here we could replace "is related to" with "is constrained by" and the meaning is obvious. In this case, where the relationship is directed, or asymmetrical, we would expect a parallel asymmetry in the data.

In keeping with the distinction between operational and conceptual terms, we speak of reciprocity instead of symmetry. Thus, if node A is linked to node B and node B is linked to node A, the A-to-B link is reciprocated. (Perhaps it is more appropriate to say "A-with-B" instead.) If A is linked to B but B is not linked to A, the A-to-B link is unreciprocated.

Clearly, we would expect reciprocated links for symmetrical relationships, and unreciprocated links for asymmetrical relationships. If our measurement techniques are reliable, this is what we will get. Many times, though, this does not happen. Why?

First, as we pointed out, the measurement techniques might not be reliable. Some relationships will be missed, leading to missing links and thus unreciprocated links. Other relationships will be wrongly identified, again leading to unreciprocated links. But this situation might not "look" any different from the second problem situation -- we may have misconceptualized the relationship as symmetrical in the first place. Maybe the relationship was really asymmetrical. This would also lead to unreciprocated links. Here, unreliable measurement would lead to some cases of reciprocation, as well as some false unreciprocated links. If the relationship was strictly asymmetrical, so that A r B implies B not-r A, all reciprocated links would be evidence of measurement error. To make matters worse, the relationship may be correctly conceptualized as symmetrical, but the operationalization may "elicit" or get at a different, asymmetrical relationship. This would lead to unexpected unreciprocated links.

Can this mess be straightened out? Only partly. In general, if our conceptualization is correct, our operationalization matched to it, and our measurement technique is reliable, the data obtained will fit the conceptual model. It is when the data do not fit the expected model that trouble should be expected. Clearly, the weakest points should be examined first and subjected to cross-validation, if that is possible.

If the measurement technique is thought to be the source of the error, and it is not possible to correct it and repeat the measurement, there are two courses of action. First, the "conservative" one: in the case of symmetrical relationships, unreciprocated links are dropped. In the case of strictly asymmetrical relationships, reciprocated links are dropped. The "liberal" approach would add the "missing halves" of unreciprocated links in the case of symmetrical relationships. In the case of strictly asymmetrical relationships, there is no "liberal" approach. Also in the case of relationships that are neither strictly asymmetrical nor strictly symmetrical, there is no way of distinguishing probable errors from correct data, and no action can be taken.

Strengths

Because the relationships are not dichotomous in the extent to which they constrain nodes, we can be more accurate when speaking of relationships if we give some idea how much constraint there is, rather than if we simply say there is or is not a relationship. This extra information will turn out to be very useful in later stages of analysis.

For the purposes of those later stages, it will be necessary to have an indicator that varies roughly as a ratio of the strength of the relationship. Because there is little ambiguity here, we call this indicator the "strength of the link." The requirement of ratio-level scaling implies that:

(1) A link from node A to node B would have a strength of zero if part A is not related to part B in the system.

(2) If the relationship from B to C constrains the included parts the same amount as the relationship from A to B, the strengths of the B-C and A-B links would be equal.

(3) If the relationship from B to C constrains the included parts twice as much as the relationship from A to B, the strength of the B-C link would be twice that of the A-B link.

Transitivity

It is not at all obvious that all relationships of the type we have been describing are transitive. However, there may be some situations in which the assumption of transitivity is made. In these cases, it will be possible to say something about links connecting nodes which are indirectly linked. For example, if A is linked to both B and C, where the relation is transitive, there should also be a link from B to C. The strength of the B-C link will be a function of not only the extent to which the relationship is transitive (this may vary) but also the symmetry of the relationship.

The transitivity issue is related to a geometric assumption often made with distance models of relational data. Many of these methods, especially in multidimensional scaling, are metric methods. The strongest assumption in these methods is also the one that gives them their power. This assumption is often called the "triangle inequality." It states that the distances between any three points must obey the law that says, "In any triangle, the length of any one side cannot be longer than the sum of the lengths of the other two sides." If network data are represented in such a way that strong relationships are replaced by short distances, the triangle inequality will say that whenever one part is related to two others, those two other parts must be related to each other. In essence, this is the assumption of transitivity.

If the relationship is only partially transitive, or if the relationship is asymmetrical, so that links are directed, the situation becomes very confusing indeed. As was mentioned earlier, the whole issue of transitivity is sticky, and needs much in the way of theoretical clarifying work. (There are some points at which transitivity would be salient in the next chapters. Because the problem has not yet been satisfactorily solved, it is usually ignored. This translates most often into a simplification of descriptions, so that both direction and strength of links are ignored for some purposes.)

In this section we have covered some of the conceptual issues regarding relationships between the parts. In the next section we present an example of an operational application of these concepts to a specific type of system -- human communication systems.

HUMAN COMMUNICATION NETWORKS

In Chapters One and Two we discussed some early conceptual models and operational methods that have been used in the study of human communication systems. In order to develop better methods to accomplish this task, we decided to first get a better understanding of what the task was about. This search for clarity led to the development of a general analytic approach, GASSICS, that can be used in any kind of system from which the appropriate data are available. At this point we turn back to a focus on communication networks in social systems.

We begin this section with a quick review of the basic system concepts, showing how they translate into communication network terms. After we have discussed the operationalization of the concepts, we turn our attention to actual data collection techniques. Here we will discuss both some general considerations and some characteristics specific to each of several methods.

General Concepts

The human communication system is an information system. This means that the relationships between the parts of the system -- the people -- are information relationships, rather than matter-energy relationships. Human behavior is regulated (constrained) by communication relationships. Social influence (constraint) is exercised through communication. Thus, the network is a representation of the real system, where nodes stand for people and links for communication relationships.

1. Content or mode

We said earlier that in systems where the interactions between the parts involve information exchanges, there could be many concurrent "modes" of interaction between the parts. In communication networks, this translates into a multiplicity of types of communication relationships. Formal communication networks in large organizations, friendship communication in social systems, new information communication in agricultural systems, communication about political events, and so on. In any given system, there could be any number of overlapping nets, with each person having a different role in each. The first thing to do is narrow down the field of interest to a single functional type of communication like, maybe, communication about matters related to getting the job done. If other areas are of interest too, these would be examined separately and then in the context provided by the whole set of relationships.

2. Strength

Strength was conceptualized as the amount of constraint or influence exerted as a result of the relationship. In the study of communication networks, strength has most frequently been operationalized as frequency of interaction. Thus, people who interact frequently are assumed to influence one another more than people who interact infrequently.

Other operationalizations add importance, so that important interactions lead to stronger links than unimportant ones. Still others measure strength in terms of the average duration of interactions, or in terms of total number of minutes spent interacting in a particular week. The important issue here is that the strength of the links has to be a single number that varies as a ratio of the strength of the relationship.

If the strength of the relationship is assumed to vary as a function of the amount of interaction, a first approximation to a ratio measure is accomplished by asking people how much time they spend talking with one another, and coding the responses in an appropriate way. For example, if we provide categories like:

(a) once a month or less;

(b) once or twice a week;

(c) once or twice a day;

(d) several times a day;

we will have to assign numbers to the categories in such a way that a ratio-level approximation is achieved. If we translate the categories into number of interactions per month, we might get:

CATEGORY		CODING
Once a month or less	=	1
Once or twice a week	=	8
Once or twice a day	=	27
Several times a day	=	64

This scheme provides a rough approximation to a ratio level scale of relationship strengths. More sophisticated methods are described in later sections.

3. Reciprocity/symmetry

The meaning of the idea of symmetry is important in communication networks, although there may often be difficulties in getting an operationalization of the relationship that produces results in agreement with the conceptualization upon which it is based. Let us explore the issue more carefully.

In many of the communication network studies that have been done using the methods described in this Part, the relationship has been "talks with." For example, "Who do you talk with about matters related to getting your job done?" or "Who do you talk with about new varieties of seed or weed spray?" or "Who do you talk with in your spare time?" The last example is clearly based on a symmetrical relationship -- there is no sense of direction or unevenness in the relationship. If one person talks to another, the second will also talk to the first in this kind of relationship, which is based on the sharing or transaction model of communication as a two-way process.

In the example about new innovations, there is a different situation entirely. The questions asked of respondents seldom take the unbiased form of "Who do you talk with ..." Rather, they will clearly ask respondents who their sources of information are. Rather than a two-way process of sharing, this is a one-way flow of information from a source to a receiver. This is obviously an asymmetrical relationship. As might be expected, the percent of links that are reciprocated is much lower in the innovation network than it is in the friendship network. Typical ballpark figures here would be about 2 percent reciprocated in the asymmetrical case and over 30 or 40 percent in the friendship one. Why aren't the figures closer to zero and 100 percent?

We discussed a number of possible reasons in the first section. There could be measurement error. The relationship could be incorrectly conceptualized. The relationship could be incorrectly operationalized. There could be a combination of all three.

The most legitimate argument seems to be that the relationship is not clearly enough understood to be correctly conceptualized and operationalized. For example, two people are seen having what appears to be a conversation. One of the two might say that it was a two-way conversation. The other may not. The relationship might seem symmetrical to the first but not the second. This kind of problem is most bothersome when the people in the system are different in terms of status, power, confidence, knowledge, and so on. The problem is that there are more than one kind of relationship, at the same time. There does not seem to be any quick way around this kind of difficulty, at least when the relationship is approached in terms of communication between people.

In other situations, the problem is clearly with the operationalization. One study, for example, dealt with "problems" in the day-to-day working of a group of enlisted men. The question asked "who they went to to get it taken care of." The investigators expected a situation completely different from the one they observed in the data. When the respondents were questioned, it was clear that they had a completely different interpretation of the question from the one the investigators thought was the only possible interpretation. Because of this mix-up, the investigators were actually measuring a relationship not at all like the one they had in mind. These are some of the difficulties encountered in the area of symmetry.

Actual Instrumentation Considerations

We have already seen a very simple example of the kind of questions that might be used to gather network data. In that example and in the discussion that followed it, it may have seemed that there was an implicit assumption that network data are gathered from people by asking them questions or giving them printed questionnaires to fill out. Although this has been the method used most often, it is not the only way. In the collection of data in rural villages in less developed areas, face-to-face interviews will have to be used instead of less personal questionnaires. Sometimes even this method will not work and the investigator will be forced to get information from "key-informants" -- people in the system who know what goes on. In other situations, other methods will have to be used. For example, a network study was done on a long-deceased seventeenth century New England village, where the data on friendship patterns was gathered from court and church records. While these other methods are the only ones that produce useful data in some situations, they will not be discussed extensively here. Instead, we will devote most of the discussion to a set of methods that is being used in studies of communication networks in large-scale organizational studies. When the differences between these fairly structured settings and the conditions of the particular system under investigation are recognized, appropriate measurement techniques should not be too difficult to develop.

In the typical organizational setting it is possible to assemble large numbers of people who can work easily with paper-and-pencil instruments. In this kind of situation, it may only take a few hours to collect the data from a system having several hundred members, if everything goes well.

There are two basic kinds of issues that have to be faced when designing instruments for use in these settings. The first concerns the translation of the conceptualized relationship into operational terms. The second concerns some practical considerations and formatting of the instrument.

1. Scaling

Once the content area has been laid out and the general format of the questions has been decided upon, the scaling of responses has to be considered. The requirement is a fairly reliable, single value that varies as a ratio of the strength of the relationship.

We discussed briefly a very simple approach to this in the beginning of this section. The problem with that simple instrument is that it ignores the difference between very important exchanges and very unimportant ones. It might be assumed that a frequent unimportant relationship is as strong as a less frequent, but more important, one. In this case, respondents might be asked to indicate how important the relationship is, in addition to how often it is used. The two numbers would then be combined into a single indicator of the strength of the relationship. The example below shows how this might be done.

> Please indicate by circling the appropriate numbers which people you talk to, how often you talk to them, and how important the interaction usually is. Use the coding system shown here.

FREQUENCY	IMPORTANCE
1 = once/month	1 = slightly important
2 = once/week	2 = moderately important
3 = once/day	3 = very important
4 = several times/day	4 = crucial to survival

NAME	FREQUENCY	IMPORTANCE
John Jones	1 2 3 4	1 2 3 4
Emily Stuart	1 2 3 4	1 2 3 4
Tony Mann	1 2 3 4	1 2 3 4
Belinda Humm	1 2 3 4	1 2 3 4
Mark Smith	1 2 3 4	1 2 3 4

It is necessary to combine the frequency and importance scales to get a single number. To do this, we would form a matrix where the rows are for the values of frequency and the columns are for importance, as shown below. We would then decide which entries have the highest and lowest values. Obviously, these would be the

		IMPORTANCE			
		Slightly	Moderately	Very	Crucial
FREQUENCY	Several/day	?	?	?	?
	Once/day	?	?	?	?
	Once/week	?	?	?	?
	Once/month	?	?	?	?

top right entry and the bottom left one in the example.

The next step is to assign the intermediate values. This is more difficult. For example, how does the top left entry compare with the bottom right one? What about other entries? If the values shown here are acceptable, the two scales can simply be multiplied together to give the final results.

In this example we formed the strength indicator by taking the product of the original scales. In other cases, we would use a linear combination instead. For example, say we had separate scales for face-to-face and telephone interactions, as shown below. We might decide that face-to-face interactions are twice as important as telephone interactions because of the additional non-verbal information that is transmitted in the face-to-face interactions. Then we would use this formula for calculating the final strength indicator: Strength = 2*Face-to-Face + Telephone.

4	4	8	12	16
3	3	6	9	12
2	2	4	6	8
1	1	2	3	4
	1	2	3	4

The Matrix with Values Filled in

Please indicate how much time you spend talking to each person in an average week (in minutes)

NAME	FACE-TO-FACE	TELEPHONE
Robert		
James		
Annie		
Frank		
Susan		

The important point here is that a single ratio-level indicator (or an approximation of one) must be available as an index of the strength of the relationship. A lot of trouble can be saved by constructing instruments so that they can be easily coded to give ratio-level data. If this is not done, the data must be transformed to give ratio data at the time of analysis, if that is possible.

2. Other instrumentation Considerations

In the discussions above we have seen several examples of instruments that might be used to collect network data. They are all variations of the same basic design. Some types seem to work better than others in different situations. For example, there are two ways of getting the respondent to provide the names of the people he or she is linked to. The first works well when there are less than about two or three hundred people in the organization. With this method, a list of all the people is provided and the respondent simply fills in the appropriate spots on the instrument. An example of this type is shown in "A" below.

How often do you interact with the people named here? Please indicate the approximate number of interactions per week for both job-related conversations and other conversations.

NAME	JOB-RELATED	OTHER
Sam		
Mary		
Bill		

A

In the column on the left, please write the names of people you talk to. In the other columns please indicate how many times you talk to these people in a typical week. Do this for both job-related conversations and other conversations.

NAME	JOB-RELATED	OTHER

B

In the first type, the respondent only has to recognize the name of the person he or she is linked to. In the second type, as shown in B, the respondent is asked to recall the names. The second type is appropriate for very large organizations, where it would be impractical to provide a list of all the names because of the length of such a list, or for systems where all the names of relevant people are not known.

There is likely to be a difference in the number of contacts reported on the two types of instruments. Specifically, since it is easier to recognize a name on a list than to recall a name from memory because the list of names serves as a prompter, there are generally more contacts reported with the first method than with the second (Farace, personal communication, November 1975).

A second way in which instruments may vary is in the method of coding the strength of interactions. A variety of approaches have been used here: (a) interaction frequency may be coded into categories as shown in "A" below; (b) interaction frequencies may be coded directly, as shown in "B"; (c) interaction duration may be coded into categories, as shown in "C"; or (d) interaction duration may be coded directly, as shown in "D".

FREQUENCY		DURATION	
1. Once/month 2. Once/week 3. Once/day 4. Several/day	How many times in the last week?	1. Less than 5 mins. 2. Less than 10 mins. 3. Less than 20 mins. 4. Less than 30 mins. 5. More than 30 mins.	How much time in the last week?
A	B	C	D

From a theoretical perspective, it would seem that the method shown in "D" above would provide the most valid information. However, it is harder to estimate durations of interactions than frequencies of interactions, as in "A" and "B", and it is harder to estimate precise numbers than simple categories, as in "A" and "C". Thus, the method shown in "A" is probably the easiest for subjects to use, while the one in "D" provides the best information. Again, there have been no empirical studies comparing the alternative methods.

When several content areas are to be used at once, it is not necessary to have a separate instrument for each one. Instead, they can be combined into a single form, with multiple columns for the different content areas. An example of this is shown below, where three separate content areas are being measured at once. In analysis, these will be treated as three separate networks which might later be compared and examined for similarities or differences.

> Please indicate how often you talk to the following people about each of the three topic areas. Use this system for coding your responses:
>
> 1 = once/month
> 2 = once/week
> 3 = once/day
> 4 = several times/day

NAME	PRODUCTION: GETTING MY JOB DONE, DAY-TO-DAY MATTERS	INNOVATION: NEW IDEAS OR WAYS OF DOING THINGS	SOCIAL RELATIONS: INFORMAL FRIENDSHIP CONVERSATIONS, ETC.
Harry			
Timothy			
Maude			
Jenny			
Donald			
Michael			

An Alternate Operationalization of Strength

Throughout the discussion of relationships we have been using "amount of communication" or "frequency of communication" as an operationalization of the "strength" of the relationship between a pair of people. The frequency with which this operational definition of relationship strength is used by investigators in the field is very high. The only significant deviation from this course that is used fairly often involves importance information as well. There have been no studies published using any other operational definition. The reason for this is probably simple inertia. People are used to these measures. These measures appear to be straightforward. They are easily "interpreted" to clients who use network analysis as diagnostic tools to monitor communication networks in their organizations.

However, these may not be the best possible measures to use for our present purposes, which are to gain a better understanding of how complex systems work. A possible alternative operationalization bypasses a step in the chain of assumptions made with the "standard" approach. In the standard approach, it is assumed that the more a pair of individuals interact, the stronger is the relationship. In the alternative approach, people would be asked directly how strong the relationship is. The instrument might look like this:

> Please indicate which of the following people influence you as you complete your day-to-day job activities. Use the following scale to show how much influence each person has on you.
>
> 0 = no influence at all
> 1 = slight influence
> 2 = moderate influence
> 3 = strong influence
> 4 = total control

An alternative scale would be:

> Use a scale of one to ten, where one is very slight influence and ten is total control.

This method bypasses the intermediate step of translating communication frequency into relationship strength. For this reason it seems to give a more direct measurement of strength. However, there are problems. First, it would seem to be more difficult for respondents to think in terms of the relatively abstract "influence" than the more concrete "How many times do you talk?" This has not been tested as yet, so the validity of this objection is in doubt. The second problem is that neither of the coding systems suggested above is known to be a ratio-level indicator. Of the two, the second would seem to be a better bet for a valid ratio-level indicator, but this has not been tested. However, the "standard" indicators have not been validated empirically either, so we are on no less firm ground with the new approach.

We are therefore faced with a trade-off: the standard approach is familiar and easily used in a consulting framework, while the new approach is more direct and more useful in a theoretical framework. Perhaps it would become more useful in a consulting framework if the clients could be educated about its benefits and interpretations. This will not happen, however, until it is empirically tested and found to be both useful and valid.

Diaries: An Alternative to Recall Methods

The "standard" instruments discussed above all ask respondents to either recall how often they talked in some past period or in a typical period. An alternative to recall techniques is the "diary" technique (Conrath, 1973). With this method, respondents carry a card on which they record each interaction after it takes place. Proponents of this method claim that it produces more accurate data than recall techniques, although this claim has never been tested empirically. In one preliminary comparison, however, in which the two methods were compared for the number of links, the difference between the two methods was less than five percent (Goldhaber, personal communication, October 1975). One important difference is this: if respondents indicate the time at which each interaction takes place, actual information flows can be monitored as they spread through the system.

Both of these points are valid and important. The reason diary techniques are not used more often is their obtrusiveness. Imagine having to carry a card around and having to write down every significant interaction. In a recent study conducted in a Canadian hospital, doctors and other staff refused to cooperate with the investigators, claiming that it wasn't possible to be bothered with the logging procedure after each interaction (Goldhaber, personal communication, 1975).

The trade-off is thus that, although the diary should in principle produce better results, it is much more obtrusive. We thus have another situation where something like the Heisenberg Uncertainty Principle in physics seems to work in a social scientific setting: the more accurate the measurement device, the more obtrusive the measurement process. At the present point it is not clear where the best combination of accuracy vs. obtrusiveness will be. Until some empirical information comparing the two methods is collected, it will not be possible to say one method is superior to the other.

The analytic method presented in the next chapter is compatible with each of the types of data, since diary data can be easily converted into the format of "standard" network data.

Higher Levels in Communication Networks

The data collection methods we have been describing all work at the same level of analysis -- the individual person. This is the appropriate level for the first time through the GASSICS loop. When the first GASSICS cycle is completed, it may be necessary to move up to the next level. In this case, nodes will be groups of people, rather than individual persons. The collection of relational data from groups is problematic.

The task of this phase of the GASSICS cycle is to determine the relationships among the groups, i.e., "Which groups are related to one another?" and "What are the characteristics of these relationships?" One approach that has been used to answer the first question is to look for links between nodes that are members of different groups. There are problems with this method.

(1) How are links between individuals who happy to be members of different groups distinguished from links between groups? If the arguments suggested by Simon (1973) are accepted, relationships between groups will not be the same as relationships between individuals.

(2) How is the strength of the relationship estimated? There are usually multiple links between the members of one group and the members of other groups. Is the strength of the link between two groups the sum of the strengths of the links between members of one and members of the other? Do multiple links imply a stronger relationship than a single link?

(3) How are indirect links treated? Do links through liaisons or others count? How are they combined with direct links?

It seems more appropriate to retreat to the conceptual definition of "link" and construct a new operational procedure for detecting links between groups than to bend the procedures used at the lower level of analysis. Thus, links are indicators of relationships. A link between groups would indicate that the groups are related to one another -- that they are constrained or influenced by the relationship. Therefore, it is not enough to say that one member of Group A is related to one member of Group B. This implies only that one member is related to the other. It says nothing about the groups.

If those two members happen to be very influential in their respective groups, the link between them might be a link between their groups after all. But this question cannot be answered with the standard type of network data that are usually gathered.

If the data are collected in a way that allows influence patterns, or perhaps information flows to be traced, it may be possible to distinguish between the links we are interested in and the ones we wish to ignore.

An alternative approach would be to go back to the system and ask the members of groups who (or which groups) they think influences their group.

It may very well be that group interactions take the form of exchanges of materials instead of exchanges of information. This would require a different type of measurement altogether.

Like the issue of transitivity, the issue of what constitutes a link between two groups is a question which needs to be answered before progress in that direction can be made. The solution to the problem will be deferred to another time.

CHAPTER SEVEN

STEP TWO OF THE GASSICS PROCEDURE: THE MEASUREMENT OF STRUCTURING

Chapter Six dealt with a method for collecting data from members of complex systems. The specific example developed there was communication systems. This chapter develops the statistics needed for the test of deviation from randomness which constitutes the second step in the GASSICS cycle.

Rather than starting right off with a set of a priori notions about structure, we are adopting a more "empirical" approach. Later in the chapter we will build upon the very simple model of structure as deviation from randomness in order to see what the model implies with regard to structure at the whole-system level. To do this we will need to use the logic of alternative descriptions to move from the microscopic level of analysis at which the model is stated to a more macroscopic level, where we can see more clearly the implications of the model on the whole system. When we have done this we will have a set of structural characteristics we can use to direct our investigations into complex systems where we have information about the interactions of the parts. In addition to seeing how we should look, we will see what to look for.

In the next section we will develop the statistical tools we need to move from the microscopic level of the raw data to the macroscopic level of system structure. There will be several steps along the way. First we will examine the idea of interdependence, applying it to communication relationships as an example. We will show how this move logically leads to the selection of a particular approach to the quantification of structure. The second step will be to work through the model for random systems. This will provide the baseline from which we measure deviations, which allows us to measure structure in the same terms we used to define it. We will base the random model on the assumption of zero constraint. In this random system there is no pattern to the set of interrelationships among the parts. Any pair of parts (nodes) is as likely to be connected as any other pair. We will take this initial statement through a transformation of alternative descriptions to a system-level statement of "random" structure. The third step will be to work through the model again, this time using observed values instead of values predicted by the assumption of randomness. The result will be a system-level description of "observed" structure. Finally, the "observed" will be compared to the "expected" -- giving a measure of the deviation from randomness -- which is how we defined structure.

Preliminary Discussion

Before we begin the operational discussion, it is appropriate to clarify a few points. First, we are dealing specifically with relationships between elements in the system. At this level of analysis, we are not interested in the microscopic issues, such as "Is there a relationship or not?" Second, in this preliminary model we found it necessary to make some simplifications. For example, in the statistics that follow, we have reduced relationships to binary all-or-none occurences. This was done at the expense of a partial loss of information, in order to keep the complexity of the statistics to a level that was manageable. In principle, however, the same ideas could be used for continuous data. There are problems, both with the use of binary versus continuous data and with transitivity assumptions that have not been worked out at this time.

Third, we have adopted a terminology that preserves the distinctions made in the preceding chapter between relationships in the system and links as indicators of those relationships, between the system itself and the network of links, and between an element in the system and a node in the network.

Although the discussion throughout the chapter is couched in terms of links between nodes which are members of networks, the fact that the same terms have been used to describe communication networks should in no way be interpreted to mean that the equations developed here are specific to social networks. The equations themselves are general, and can, in principle, be used at any level of analysis in any system for which the proper relational data are available.

Our final goal is a set of measures with which we may operationally determine how much structuring there is at the whole-system level. After we have derived the appropriate equations, we will show by use of some examples how this measure is related to the amount of organization in a few simple systems. In a discussion at the end of the section we will suggest some implications this approach has for the question we asked earlier: "What should we look for?"

OPERATIONALIZATION

Let us begin with a system composed of N elements with L links distributed among the elements in some way. Each link indicates that a certain pair of elements is connected by some kind of functional relationship, the exact nature of which is not immediately important here. Let us assume that the number of links, L, is less than the total number possible, which is $L_{max} = \frac{N(N-1)}{2}$. This implies that nodes, on the average, will have less than the maximum possible number of links to other nodes that they could have, i.e., the system's density (or connectiveness), expressed as the ratio of the observed number of links divided by the maximum possible, is less than 1.0.

$$\text{System Density (connectiveness)} = \frac{2L}{N(N-1)} \qquad \text{(Eq. \#1)}$$

If the links are randomly assigned to nodes, without duplication or reflexive links, the average number of links per node will be $\frac{2L}{N}$. Since there are always exactly N-1 possible alternatives for each node to have links with, there is some freedom in terms of where the links will actually be. For example, if there are 11 nodes and 22 links, there will be, on the average, four links for each node.

(It takes two nodes for each link.) For each node there are 10 other nodes that could be connected to with links.

Now, if our network is truly random, we will have no clue as to where those links will be. They will distribute normally across the entire network. Barnett (1973) showed that in such a network there is no differentiation of the system into parts. The relationships between individual elements will be independent of one another. That is, even if we know where some of the links are — if, for example, we know that B is linked to A and C -- we can say nothing about where the other links are; in other words, we cannot tell if A is linked to C. This example is shown in Figure 1.

On the other hand, if the network is structured, the relationships between pairs of nodes will not be independent of one another. In other words, in a structured network, the links will be at least partially interdependent. We would like to be able to measure the amount of interdependence. But how can we do this, when the only information we have is who is linked to who?

The logic of alternative descriptions in complex systems tells us that if we focus our attention on individual nodes or links we will not be able to see the kind of interdependence we are interested in here. We need some way of looking at larger sets than individuals, since interdependence requires more than one unit. In fact, since we are interested in the interdependence of links, we need to look at sets that are large enough to include at least two links. The smallest set of nodes that will include two links has three members. But how does this help us? How can we tell if any pair of links are interdependent? With the data we have, we cannot. If all we know is that A is linked to B, we can say nothing about possible links from A or B to some other node, say C. C could be any other node in the system. There is no reason why we would or would not expect a link to C.

However, if we expand our set to include three links, we do much better. If both A and B are linked to C, we can look at the possibility of a link from A to B. If A, B, and C all have to coordinate their activities for some reason, we would expect a link from A

Figure 1

CONSTRAINT AND THE ORGANIZATION OF LINKS

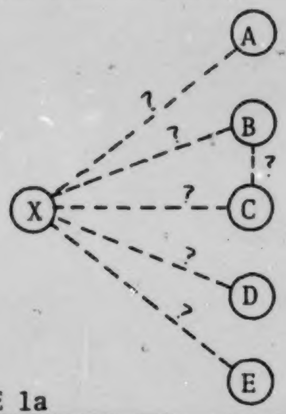


FIGURE 1a

If we expect X to have some number of links, say 2, but know there is no constraint, the links could be anywhere. Furthermore, we have no clues about links between other nodes to which X might be linked -- for example, B and C.

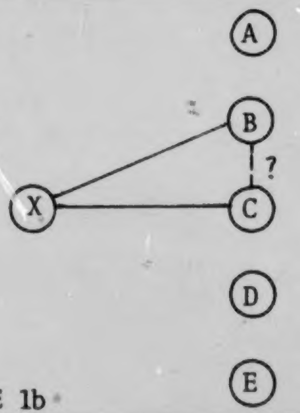


FIGURE 1b

Even if we know that X is linked to B and C, we can only guess about whether or not B is linked to C, since the links are all independent of one another.

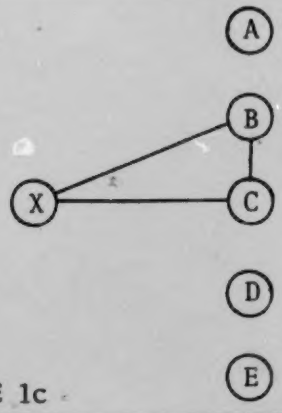


FIGURE 1c

If the links are not independent, that is, if they fit together in some sort of pattern, we are in a better position to predict what will happen. In this example, we might find that if X is linked to both B and C, the probability of B having a link with C is twice as high as it would be if X did not have those links.

to B. If, however, A and B are independent of one another, we would not expect them to be linked any more than the probabilities would predict for any pair of randomly selected nodes. Here we have a way of telling how much interdependence there is — by counting the number of occurences of sets of three nodes with three links. We can calculate the number of these "triangles" we would expect by chance alone (that is, in a random network), and compare this "expected" number to the number we actually observe. If we observe more or less than expected, we can infer the existence of some structuring "force" which operates at the level of the whole system. Why else would people organize themselves in such a way as to maximize or minimize the number of triangles?

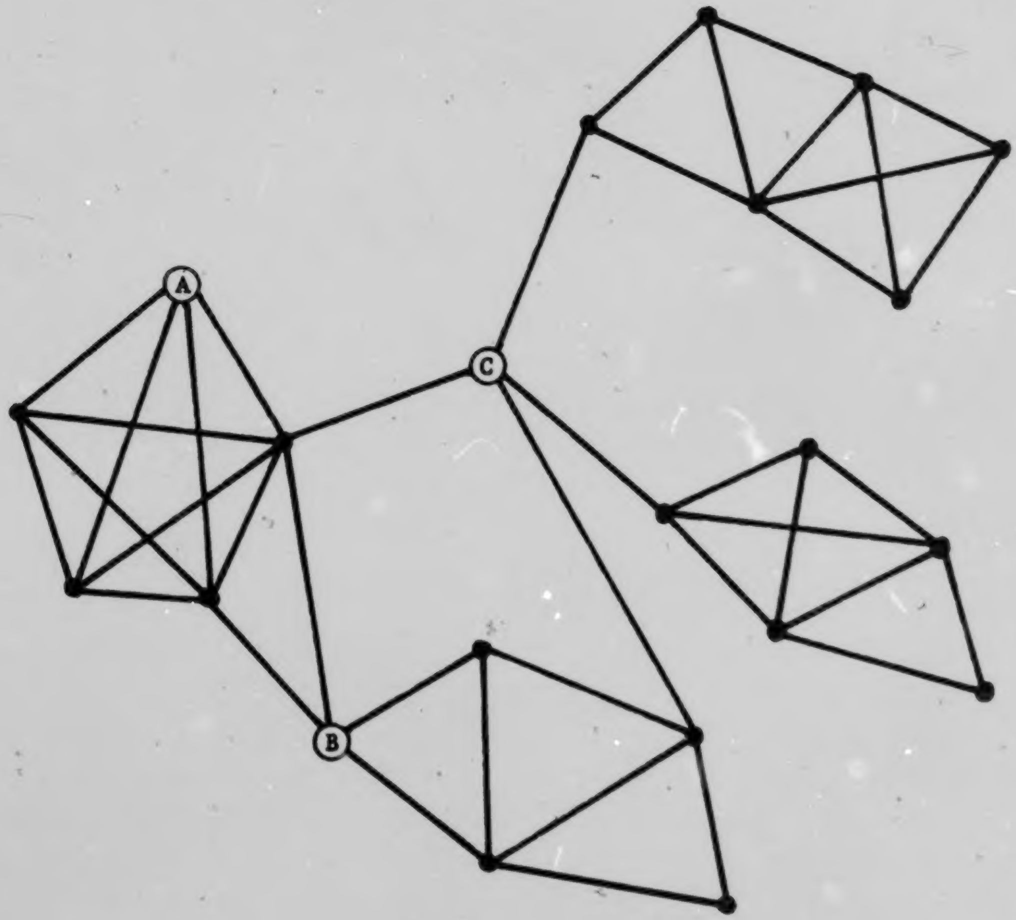
"Triangles" and Constraint

Let us explore the concept of "triangles" a bit more before we take up the statistics of the situation. It is probably not yet clear why the number of triangles is related to structuring. Let us see what the nodes in a network have to do in order to "make" triangles. In order for there to be a triangle, a node must interact with other nodes which themselves interact with each other. If the original node interacts with only a few of these tightly interconnected nodes, and "wastes" the rest of its links interacting with nodes that are not linked to the tightly connected group, the number of triangles there can be is reduced. This is shown in Figure 2, where A has the maximum number of triangles that is possible, given that it has four links. Node B, however, has divided its links among two sets of nodes which themselves are not connected. As a result of this, it has only two triangles -- one-third of the maximum for a node with four links. Node C is an even more extreme case -- here the links are spread in such a way that there are no triangles whatsoever.

The point to be understood from all this is that in order to maximize the number of triangles, links must be between nodes which already have a large number of mutual contacts. There will be groups, each of which is composed of a set of nodes, which, to a large extent,

Figure 2

INTERCONNECTION AND THE NUMBER OF TRIANGLES



confine their interactions to other members of the same group. There will be very few links between these groups, because this would lower the number of triangles that could be formed. The groups will be arranged in a very peculiar way — in general, the number of members in each group will be equal to roughly the number of links had by each member of the group. In addition, all the nodes in each group will tend to have the same number or very close to the same number of links. In other words, each member in a group having eight members will be likely to have seven links. Nodes with three links will be grouped in sets of four, and so on. An example of such a network is shown in Figure 3.

The highly structured network shown in Figure 3 can be contrasted with one in which the number of triangles is minimized. In this kind of situation, nodes do not have links with other nodes which are linked to each other. There will be no differentiation into groups because this is what raises the number of triangles. An example of this situation is shown in Figure 4.

The following exercise will help to make the preceding discussion clear.

> Construct a network of about 10 or 12 nodes. Put in enough links to give a connectiveness of about 0.25. Use a random number table to assign the links to pairs of nodes, by drawing pairs of random numbers and connecting the nodes indicated by the numbers. Don't duplicate any links and don't link any nodes to themselves. Count a link from A to B the same as one from B to A. Examine the overall network obtained. It is not likely to show any patterning or differentiation into groups. Now, repeat the process, except add the rule that allows only even-numbered nodes to have links to other even-numbered nodes, and only odd-numbered nodes to have links to other odd-numbered nodes. Now, count the number of triangles in each network. There should be more in the second than in the first. If the process is repeated again, dividing the entire set into three or four subgroups, the results are even more striking.

Figure 3

**THE MAXIMIZING NUMBER OF TRIANGLES
WITH GROUP STRUCTURE**

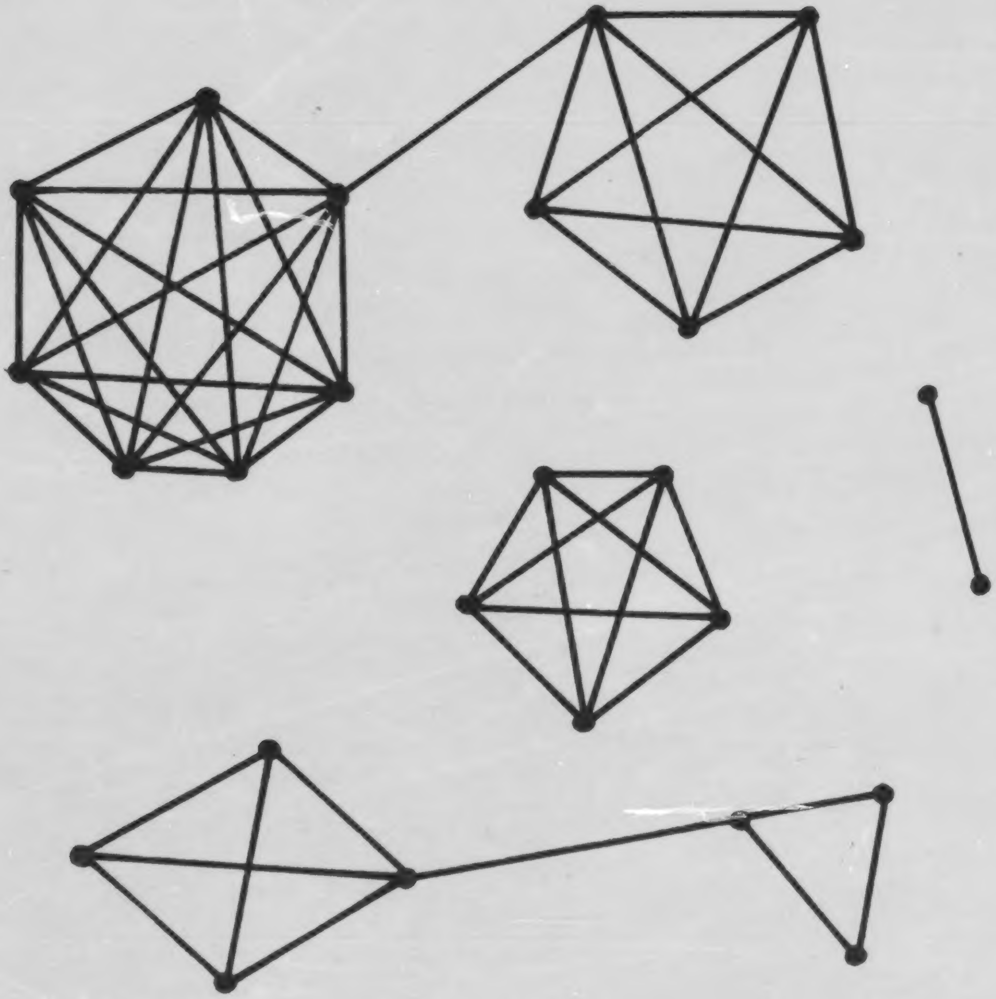
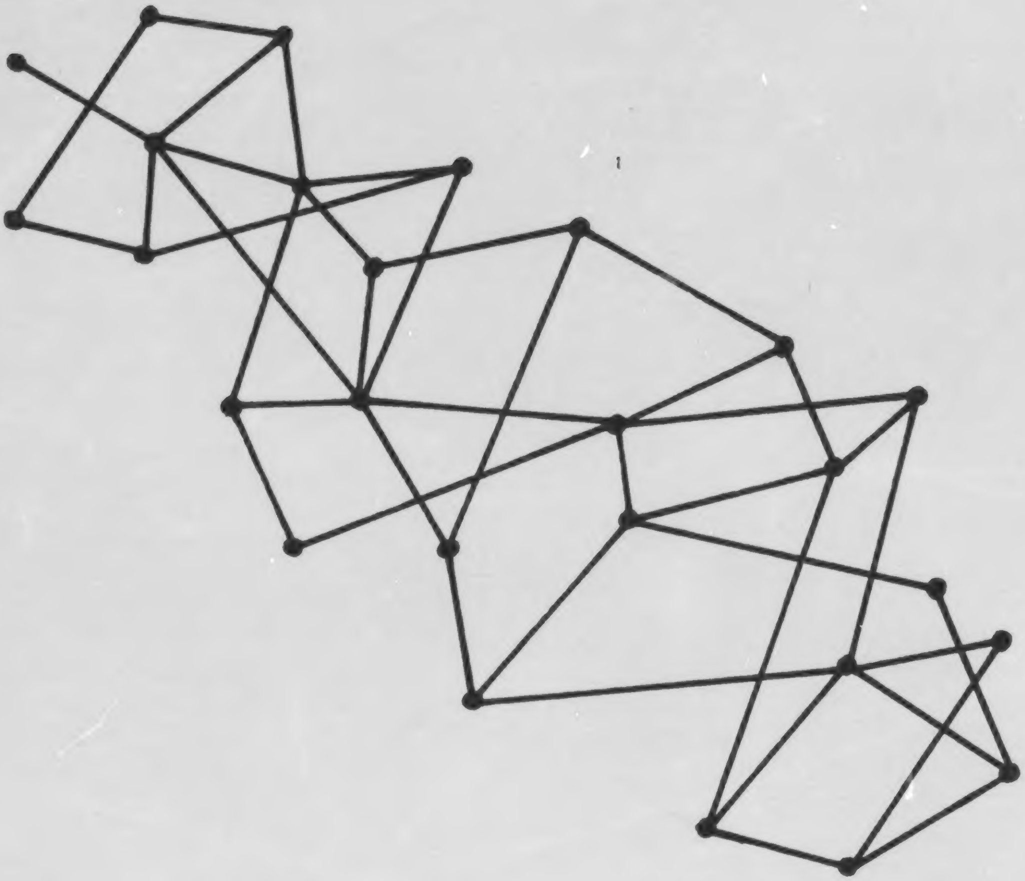


Figure 4

MINIMIZING THE NUMBER OF TRIANGLES
WITH GROUP STRUCTURE



The more constraints that are imposed on the set of interactions among the members of the system, the more triangles there are, until we reach a maximum where further increases in constraint are no longer possible.

Thus, the number of triangles is directly related to the amount of constraint and, therefore, the amount of structuring in the network. All of these three are proportional to the amount of differentiation of the system.

Obviously, the number of triangles is a very powerful descriptor of the network. But this is only true if all networks are the same size and have the same number of links. It remains to put this measure into the form of a standardized metric, by expressing it as a fraction of the maximum number of triangles possible, given the size and density of the network. In addition, we will need to be able to calculate the number of triangles we would expect in a random network, for use as the baseline from which we will measure deviations.

The Expected Number of Triangles: The Random Model

The derivation for the expected number of triangles in a random network is as follows.

Given any link, say from A to B, there are N-2 possible triangles involving this original link because there are N-2 other nodes which are all candidates for the third vertex of the triangle. For any particular chosen pair of nodes, the probability of there being a link connecting the pair is $\frac{2L}{N(N-1)}$, which equals the system density. There are L "original" links, and if we eliminate duplications involving permutations of order we get a total expected number of triangles of:

$$T_e = (N-2)\left(\frac{2L}{N(N-1)}\right)^2\left(\frac{L}{3}\right) = \frac{4\,L^3(N-2)}{3(N(N-1))^2}, \qquad \text{(Eq. \#2a)}$$

or, in parametric form, using system density, C, as a scale factor:

$$T_e = (N-2)\left(\frac{L}{3}\right)C^2 \qquad \text{(Eq. \#2b)}$$

This number will be seen to be equal to the maximum number of triangles possible whenever L is at its maximum of L_{max}. This maximum number of triangles is given by:

$$T_{max} = \frac{N(N-1)(N-2)}{6} \qquad \text{(Eq. \#3)}$$

In addition, the ratio between T_e and T_{max} is related to system density, C, as:

$$\frac{T_e}{T_{max}} = \frac{(N-2)(\frac{L}{3})\ C^2}{\frac{N(N-1)(N-2)}{6}} = \frac{2C^2L}{N(N-1)} = C^3 \qquad \text{(Eq. \#4)}$$

Variance in the Distribution of Links

These calculations were based on only two parameters -- L, the number of links, and N, the number of nodes. If the network is totally random, the links will distribute normally across the entire network. This means that we could examine the distribution of links across nodes and would expect that this distribution would also be normal. Let ℓ_i be the number of links with node i. In general, the ℓ_i's must satisfy the equation $\Sigma\ell_i = 2L$, where L is the total number of (bidirectional) links. Also, no ℓ_i may be greater than N-1 or less than zero. Since there are N nodes, the mean of the ℓ_i's, $\overline{\ell}$, must be $\frac{2L}{N}$. It, too, has a maximum value of N-1.

In the totally random network, the probability of any particular link is $\frac{2L}{N(N-1)}$. This can be derived either by dividing the number of links observed by the total number possible, or by dividing $\overline{\ell}$ by the maximum number for each node.

Since each node has N-1 chances to have links, with $p = \frac{2L}{N(N-1)} = C$, the binomial expansion gives the mean, $\overline{\ell}$, of $np = \frac{2L(N-1)}{N(N-1)} = \frac{2L}{N}$, which agrees with the value derived above. The expected variance would be $npq = S_e^2 = np(1-p)$, again by the binomial expansion. This is:

$$S_e^2 = (N-1)(\frac{2L}{N(N-1)})(\frac{N(N-1)-2L}{N(F-1)}) = \frac{2L(N(N-1)-2L)}{N^2(N-1)} \qquad \text{(Eq. \#5)}$$

Figure 5a shows a plot of the expected variance against L, for different values of N. Figure 5b shows a general plot of S_e^2 against L and N.

The F test can be used to compare an observed S^2 to the expected value of S_e^2. The observed value, S_o^2, is calculated as:

$$S_o^2 = \frac{1}{N}\Sigma(\ell_i - \bar{\ell})^2 \quad \text{or} \quad \frac{\Sigma\ell_i{}^2 - \frac{1}{N}(\Sigma\ell_i)^2}{N} \qquad \text{(Eq. \#6)}$$

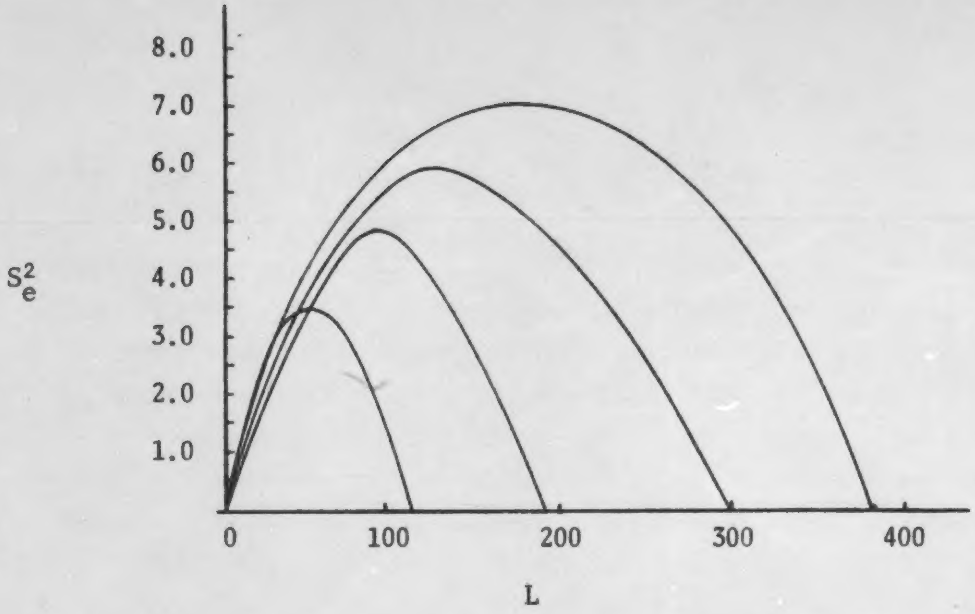
The degrees of freedom are N-1 and N-1. Figure 6 shows a general plot of S_e^2 against C and N with significance regions for $p \leq .05$ and $p \leq .01$.

A significant deviation from S_e^2 means that the distribution of ℓ_i's differs from what is expected by chance. This would mean that either some nodes have a disproportionately large number of the links, compared to the rest of the nodes, or that the nodes have unexpectedly uniform numbers of links, compared to what is expected by chance. In general, S_o^2's higher than S_e^2 will lead to inflated measures of structure, while values lower than S_e^2 will tend to be biased in the opposite direction. This is illustrated in Figure 7. In these cases, we would use measures for T_e which have been corrected for the particular distribution of ℓ_i's. These measures are discussed in the following paragraphs.

Correcting T_e for the Effects of ℓ_i

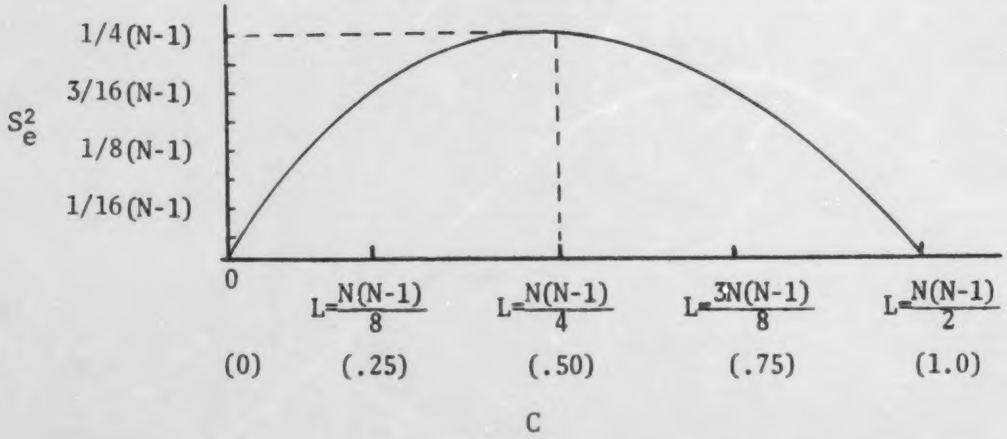
What if the distribution of ℓ_i's is not normal? For example, most of the links could be concentrated in a small part of the network. We would expect a different value of T_e if this were the case. It is possible to "partial out" the effect of different ℓ_i's by taking the following approach. For every triangle it is a part of, a node must have a pair of links. For a node with ℓ links, there are $\frac{\ell(\ell-1)}{2}$ such pairs. Summing across all nodes, we get the total number of these pairs: Total pairs $= \frac{1}{2}\Sigma(\ell_i(\ell_i - 1))$. Each of these pairs defines a triangle when the third link is added. The probability of this third link is equal to the probability of a link between any given pair of nodes, or $\frac{2L}{N(N-1)}$. Dividing by three to eliminate duplication caused by order permutations, we get:

FIGURE 5a



Plot of expected variance against number of links, for N's (number of nodes) of 16, 20, 25, and 28.

FIGURE 5b



Generalized plot of expected variance against C. Note the maximum value for S_e^2 of 1/4(N-1) at the point where system density = C = 0.50 ($L=1/2\ \frac{N(N-1)}{2}$), or half the maximum number of links.

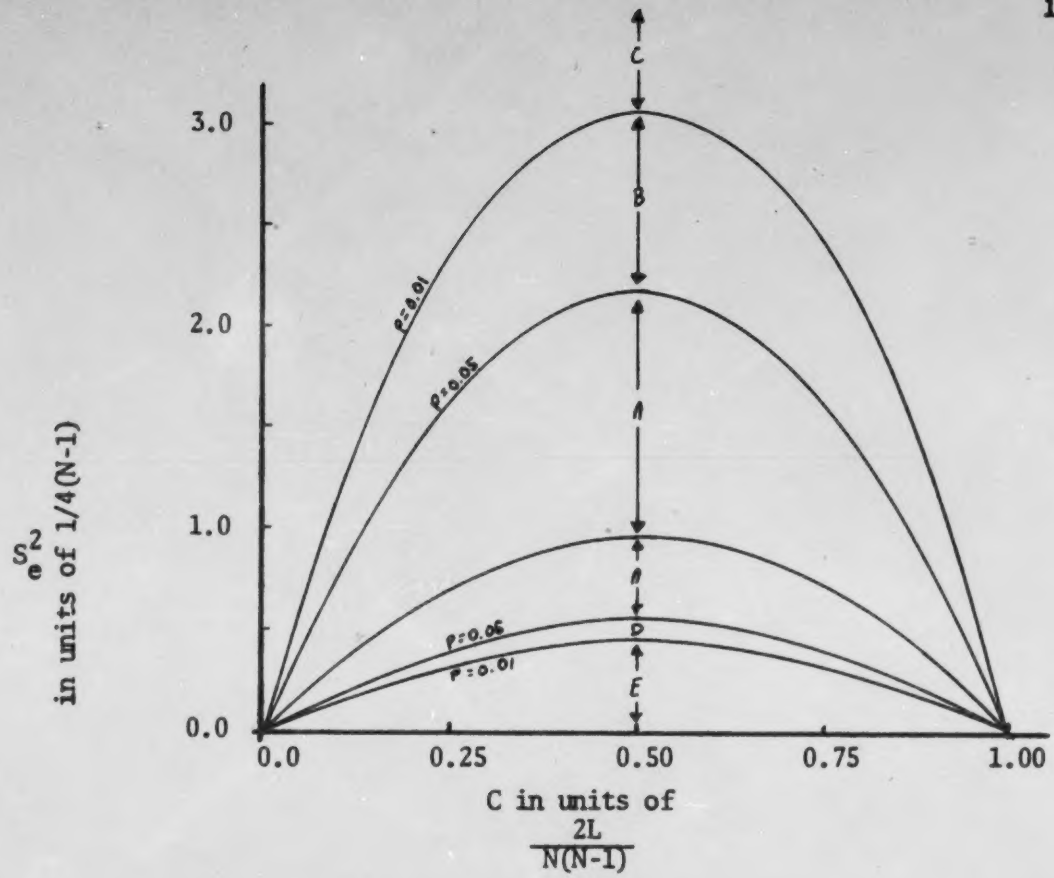


FIGURE 6

Generalized plot of S^2 in units of $\frac{N-1}{4}$(ordinate) against system density, C, in units of the observed number of links, L, expressed as proportion of the maximum of $\frac{N(N-1)}{2}$ (abcissa). (Significance values for N=20, df=(19,19)).

Region A - Values of S_o^2 in this region are not significantly different from the expected value, $p = 0.05$.

Region B - Values of S_o^2 in this region are significantly higher than expected. $p<0.05$.

Region C - Values of S_o^2 in this region are significantly higher than expected. $p<0.01$.

Region D - Values of S_o^2 in this region are significantly lower than expected. $p<0.05$.

Region E - Values of S_o^2 in this region are significantly lower than expected. $p<0.01$.

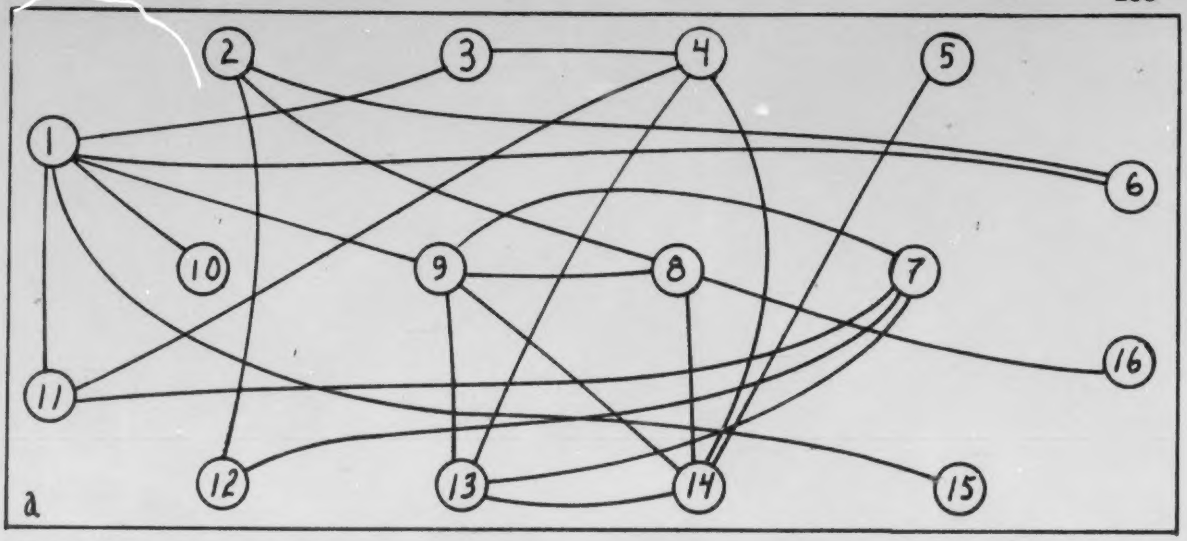

a
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 7a

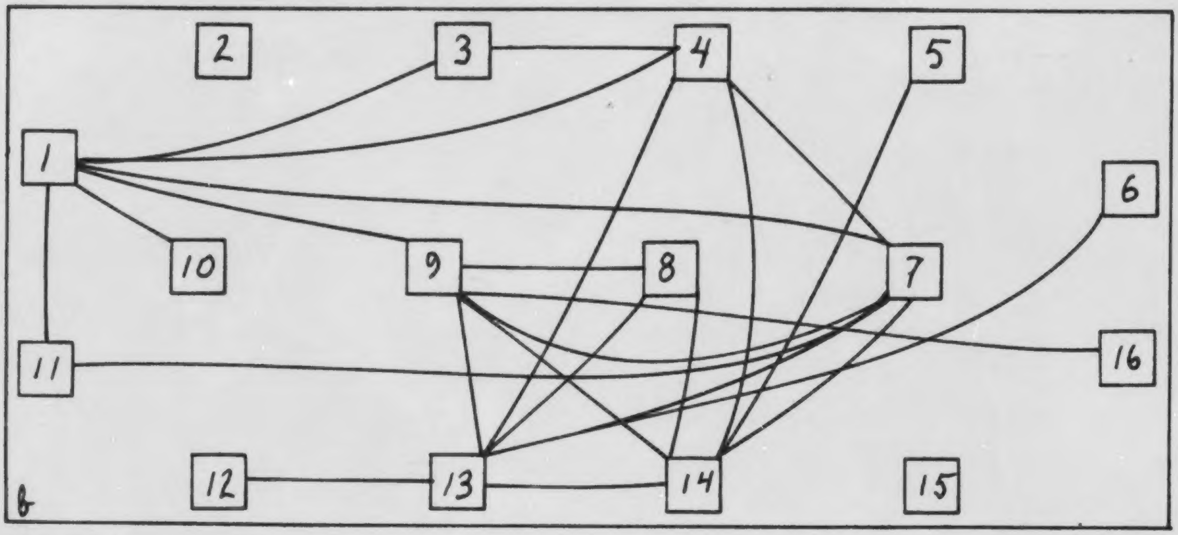

b
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 7b

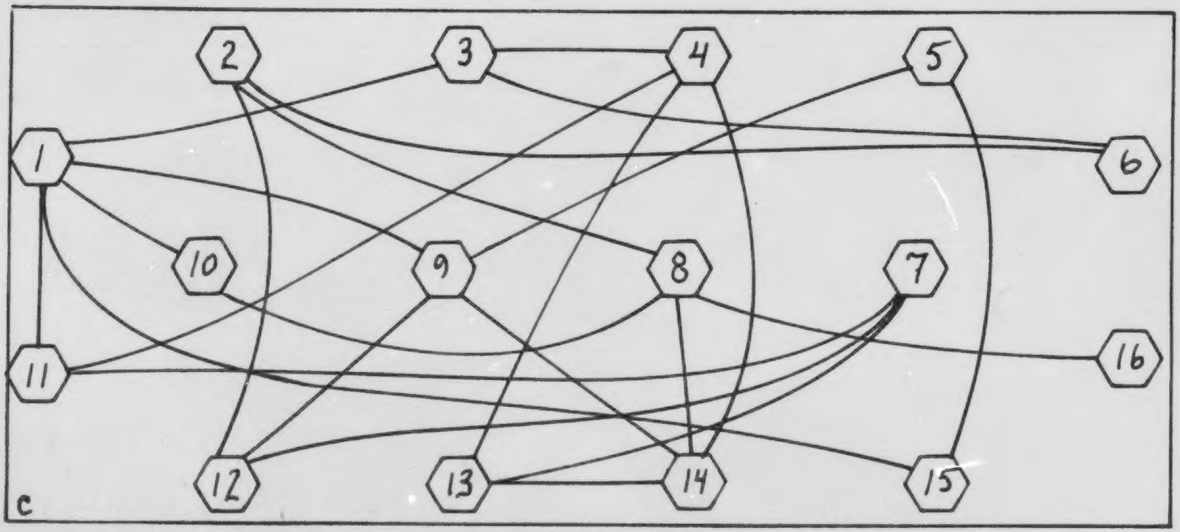

c
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 7c

Figure 7

THE RELATION BETWEEN S_o^2 AND THE EXPECTED NUMBER OF TRIANGLES

Table 1a

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\Sigma \ell_i$
ℓ_i	a	6	3	2	4	1	2	4	4	5	1	3	2	4	5	1	1	48
	b	6	0	2	5	1	1	6	4	5	1	2	1	7	6	0	1	48
	c	5	3	3	4	2	2	3	4	4	2	3	3	3	4	2	1	48

Table 1b

	$\Sigma \ell_i$	$\bar{\ell}$	S_o^2	S_e^2	T_o	T_e	$T_{e.\ell}$
a	48	3	2.5	2.4	4	4.4	4.5
b	48	3	5.75	2.4	14	4.4	6.25
c	48	3	1.0	2.4	1	4.4	3.8

In [a] is shown a network, constructed by drawing pairs of numbers from a random number table. The ℓ_i's distribute approximately normally, with S_o^2 close to S_e^2. In this case, the expected number of triangles, T_e, was close to the observed number of 4.

The network shown in [b] was constructed from the one used in [a] by rearranging the links so that S_o^2 would be increased. Thus, S_o^2 does not equal S_e^2. In this case, the expected number of triangles, T_e, underestimates the true value of $T_o = 14$.

The network shown in [c] was constructed from the one used in [a] in such a way as to reduce S_o^2. In this case, T_o, 1, was overestimated by T_e.

The relevant values are tabled above for easy comparison.

$$T_{e.\ell} = \frac{L\Sigma\ell_i(\ell_i - 1)}{3N(N-1)} \qquad \text{(Eq. \#7)}$$

for the expected number of triangles, after removing the effects of ℓ_i. This value will approximate T_e when ℓ_i distributes normally with $S_o^2 = \frac{\Sigma(\ell_i - \ell)^2}{N}$, the observed variance, close to the expected value of $S_e^2 = \frac{2L(N(N-1)-2L)}{N(N-1)}$.

Maximum Limits to T and S^2

If L, the number of links, is some number less than the maximum number possible, L_{max}, the number of triangles can be maximized if the links are constrained to a subset of the total population of N nodes. This subset will maximize the number of triangles when its size, n, satisfies this equation: $L = \frac{n(n-1)}{2}$. This value is approximated by $n = \sqrt{2L}$, for large L. The equation for the maximum number of triangles for a network with L links then becomes:

$$T_m = \frac{\sqrt{2L}(\sqrt{2L} - 1)(\sqrt{2L} - 2)}{6} \quad \text{or} \quad \frac{n(n-1)(n-2)}{6} \qquad \text{(Eq. \#8)}$$

If the system density, C, is compared to the ratio between T_{max}, the total maximum possible number of triangles for a network of size N with L_{max} links -- $\frac{N(N-1)(N-2)}{6}$, and T_m, the following relationship is observed:

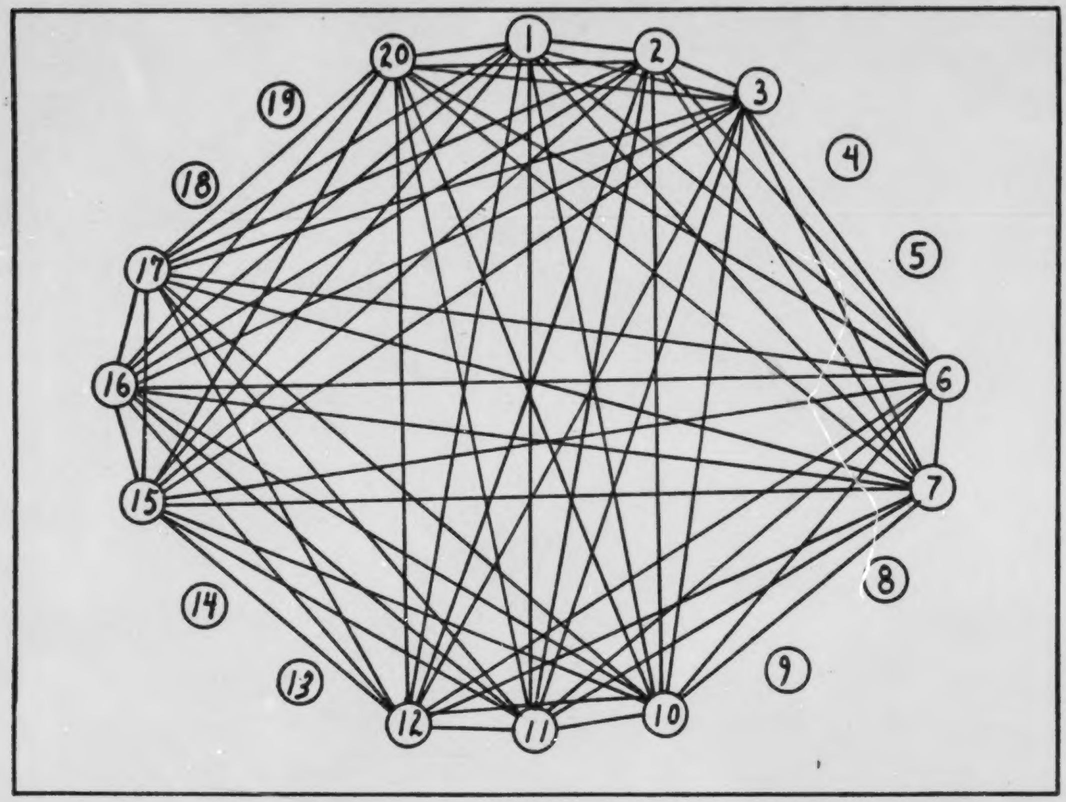
$$C^{1.58} \approx \left(\frac{L_o}{L_{max}}\right)^{1.58} \approx \left(\frac{T_m}{T_{max}}\right) \quad \text{or} \quad \left(\frac{2L_o}{N(N-1)}\right)^{1.58} \approx \frac{n(n-1)(n-2)}{N(N-1)(N-2)} \approx C^{1.58}$$

This situation is illustrated in Figure 8, where all the links are constrained to a subset of the total population of nodes. Obviously, this is a trivial maximum. Notice that under these conditions, S_o^2 for this particular N and L will be given by the equation:

$$S_n^2 = \frac{\Sigma(\ell i-\bar{\ell})^2}{N} = \frac{n((n-1)-\bar{\ell})^2 + (N-n)\bar{\ell}^2}{N} \qquad \text{(Eq. \#9)}$$

Figure 8

MAXIMIZING THE NUMBER OF TRIANGLES
BY REDUCING THE EFFECTIVE N



N = Number of Nodes = 20

L_o = Number of Links Observed = 66

$n = \sqrt{2N} = 12$

$$L_{max} = \frac{N(N-1)}{2} = 190$$

$$T_{max} = \frac{N(N-1)(N-2)}{6} = 1140$$

$$T_m = \frac{n(n-1)(n-2)}{6} = 220$$

$$\frac{L_o}{L_{max}} = \frac{66}{190} = .346$$

$$\frac{T_m}{T_{max}} = \frac{220}{1140} = .193$$

$$(.346)^{1.58} = .193$$

There are twenty nodes, but the links are confined to a subset of only twelve. The ratio of observed links to the maximum possible is .346. When raised to the 1.58 power, this value is very close to the value of the ratio of maximum number of triangles for any network having 66 links to the maximum number of triangles for any network having twenty nodes.

where N is the number of nodes, n is approximately $\sqrt{2L}$, and $\bar{\ell}$ is the mean of the ℓ_i's.

The maximum number of triangles for a given distribution of ℓ_i's can also be calculated, giving an upper limit of $T_{m.\ell}$ in which the effects of the ℓ_i's have been "partialled out" as they were for $T_{e.\ell}$. Let the number of links with node i be ℓ_i. For a node to be part of a triangle it must have two links, forming two of the three edges. For each pair of links with a given node, there is a single possible triangle. The maximum number of triangles for any single node with ℓ_i links will then be equal to the number of pairs of links with that node, or $\frac{\ell_i(\ell_i - 1)}{2}$.

If we sum across all nodes and divide by three to eliminate duplication, we get:

$$T_{m.\ell} = \frac{1}{6}\Sigma\ell_i(\ell_i - 1) \qquad \text{(Eq. \#10)}$$

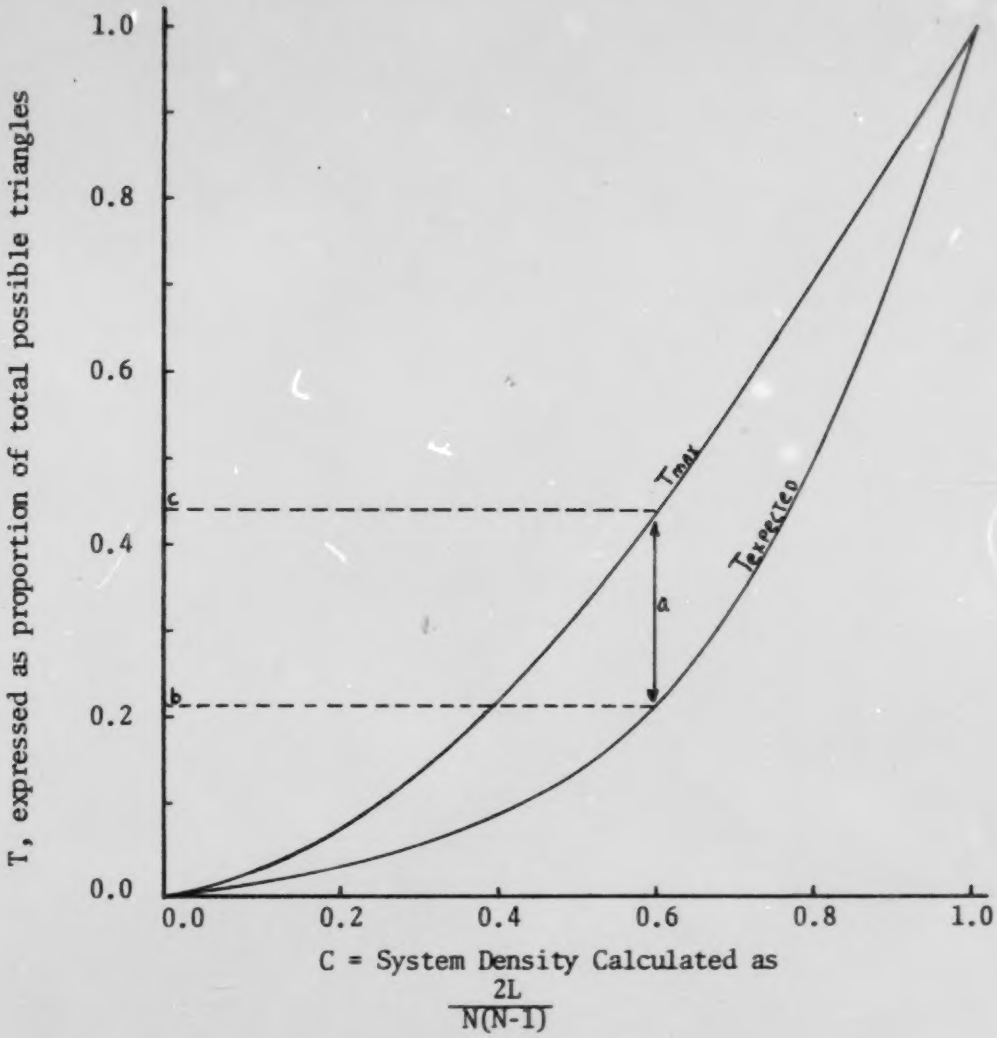
which is the maximum possible number of triangles, given the set of N nodes and their ℓ_i's. This number will equal T_m, T_e, and $T_{e.\ell}$ when every node has the maximum number of links possible, which is N-1. When the total number of links is less than L_{max}, $T_{m.\ell}$ will be less than T_m, but it will never be less than T_e or $T_{e.\ell}$.

Relative Structure Measures

We have defined upper limits for the number of triangles, given the number of links per node, and the number of nodes. These limits are reached only under conditions of maximum constraint. We have also shown how to compute the expected number of triangles, given conditions of zero constraint and how to remove the effects of a particular set of ℓ_i's from the expected number of triangles. We have thus defined a range over which the number of triangles can vary as constraint varies. This range is shown graphically in Figure 9.

Now, if we count the number of triangles in any given network, subtract the number expected by chance, and divide the result by the difference between the maximum possible and the number expected, we will obtain a value that ranges from 0.0 to 1.0; where zero indicates

FIGURE 9



Generalized plot of maximum number of triangles, T_{max}, and expected number of triangles, T_e, against C, system density.

For a network with 100 nodes, the maximum number of links would be 1/2(N(N-1)) = $\frac{100\ 99}{2}$ = 4950. If the observed number of links is 2970, the proportion of links would be $\frac{2970}{4950}$ = 0.6. With this density, the range of triangles would lie on the arrow marked "a". The expected proportion of triangles is given by $(.6)^3 = .216$, which is point "b" on the ordinate. The maximum number of triangles as a proportion of the total possible for this N is given by $(.6)^{1.58} = .45$, which is "c" on the ordinate. The maximum number of triangles for N = 100 is given by $\frac{(100)(99)(98)}{6} = 161700$. Thus, $T_e = (.216)(161700) = 34700$.

no structure other than that expected by chance and one indicates the amount expected under conditions of maximum possible constraint. This measure is insensitive to number of links, L, or number of nodes, N, and thus can be used to compare networks of various sizes and densities. The final equations are given here:

$$S = \frac{T_o - T_e}{T_m - T_e} \qquad \text{(Eq. \#11)}$$

where T_o is the observed number of triangles, T_e is the expected number, $\frac{4L^3(N-2)}{3(N(N-1))}$; and T_m is the maximum, $\frac{\sqrt{2L}(\sqrt{2L} - 1)(\sqrt{2L} - 2)}{6}$.

Controlling for the effects of ℓ_i, we get:

$$S_{.\ell} = \frac{T_o - T_{e.\ell}}{T_{m.\ell} - T_{e.\ell}} \qquad \text{(Eq. \#12)}$$

where T_o is the observed number of triangles; $T_{e.\ell}$ is the expected, controlling for ℓ_i, $\frac{L\Sigma\ell_i(\ell_i - 1)}{3N(N-1)}$; and $T_{m.\ell}$ is the maximum, controlling for ℓ_i, $\frac{1}{6}\Sigma\ell_i(\ell_i - 1)$.

SYSTEM-LEVEL IMPLICATIONS OF STRUCTURING

How is a structured network different from a random or non-structured one? A structured network will be more organized -- its members may be arranged into groups or some other type of orderly pattern. The more structured the network is, the more highly differentiated it will be. The networks shown in Figure 10 will help to make this clear. In Figure 10a, we have a random network -- the system as a whole does not break down into parts; there are no intermediate-level groups composed of individuals. Because of the low number of triangles, there is relatively little opportunity for the individual nodes to work together to coordinate their activity. In order for a number of nodes to work together as a group, they must be able to interact with each other enough to coordinate their behavior; this seems to demand network configurations rich in triangles.

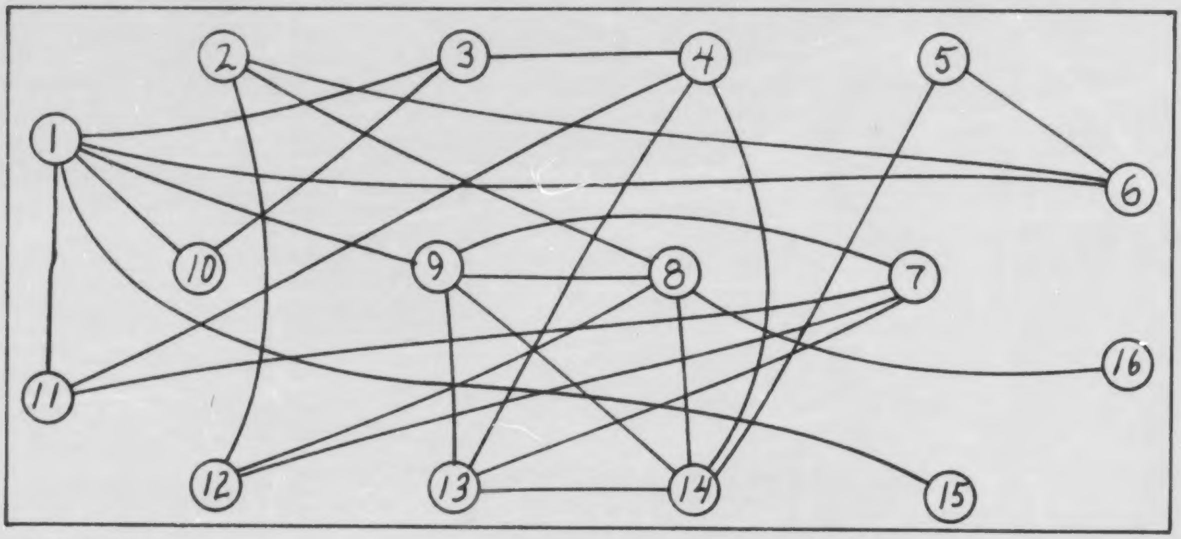

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 10a

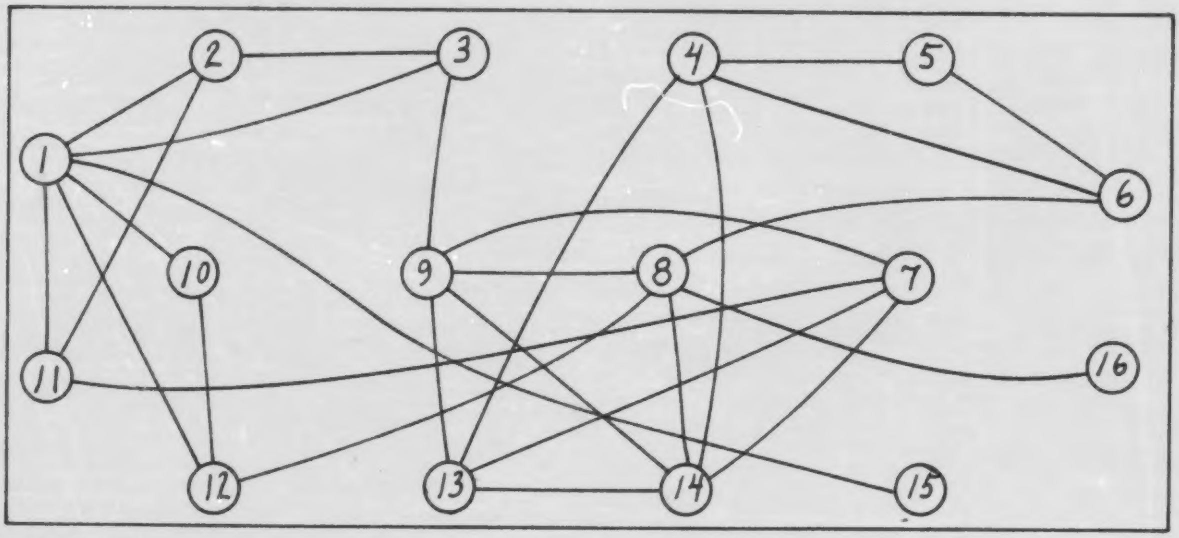

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 10b

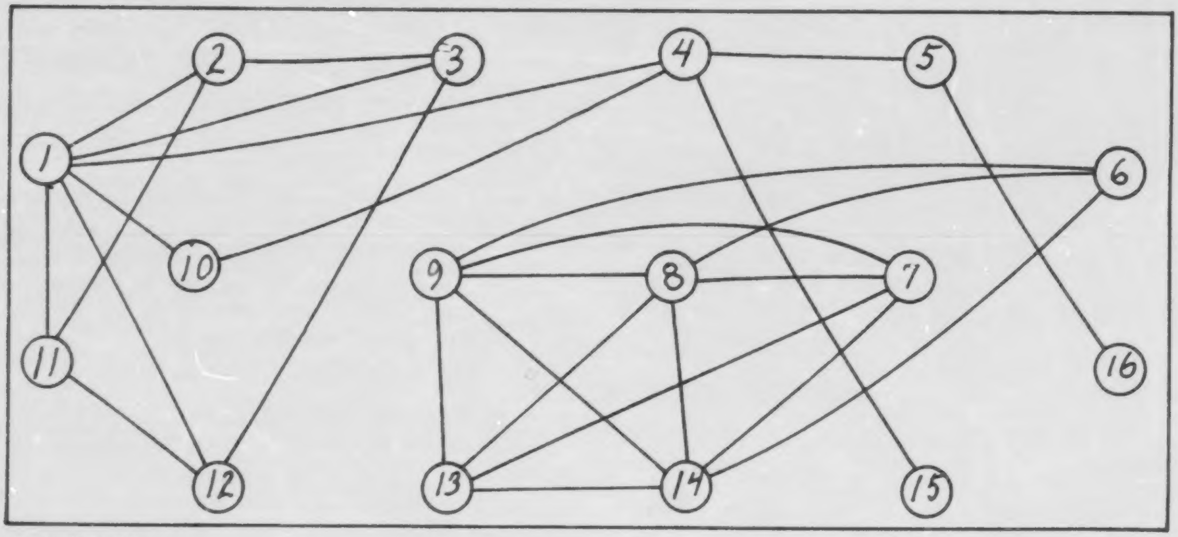

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

FIGURE 10c

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
ℓ_i	6	3	3	4	2	3	4	5	5	2	3	3	4	5	1	1

TABLE 2a

The distribution of ℓ_i's for the networks shown in Figure 10. The distributions are identical for all three networks, and are approximately normal. S_o^2 is close to S_e^2.

	N	L	$\bar{\ell}$	S_o^2	S_e^2	T_o	T_e	$T_{e.\ell}$	T_m	$T_{m.\ell}$	S	$S_{.\ell}$
A	16	27	1.69	1.97	1.85	6	2	6	35	26	.138	.00
B	16	27	1.69	1.97	1.85	9	2	6	35	26	.241	.15
C	16	27	1.69	1.97	1.85	17	2	6	35	26	.517	.55

TABLE 2b

In this table are shown relevant values describing the three networks shown in Figure 10. Values for T_e and $T_{e.\ell}$ were rounded to integer values. Lists of the triangles for the three networks appear below.

NETWORK A

1-3-10
2-8-12
4-13-14
7-9-13
8-9-14
9-13-14

NETWORK B

1-2-3
1-2-11
1-10-12
4-5-6
4-13-14
7-9-13
7-9-14
8-9-14
9-13-14

NETWORK C

1-2-3
1-2-11
1-4-10
1-11-12
6-8-9
6-8-14
6-9-14
7-8-9
7-8-13
7-9-13
7-9-14
7-13-14
8-9-13
8-9-14
8-13-14
9-13-14

The networks shown in Figures 10b and 10c were constructed from this first network by rearranging the links so that there would be more triangles in b than in a, and even more in c. In both b and c, the ℓ_i's are exactly the same as in a. All the relevant values are shown in Table 2.

Even a quick glance will show that the trend is clear as we move from the random network shown in Figure 10a to the structured one in Figure 10c — the system shown in c exhibits clear differentiation into parts. This is a direct result of the increase in structure, as operationalized by the number of triangles. The relation between number of triangles and system differentiation is due to the fact that in order to have groups, we must have sets of nodes which interact more with each other than with other nodes outside the sets. This would lead to an abundance of cases in which the nodes linked to any particular node are linked to each other, which is exactly what triangles are.

This, then, suggests at least a partial answer to our question, "What do we look for?" We look for structure in terms of differentiated parts within the system. These differentiated parts -- call them "groups" -- will be made up of individual elements which have most of their interactions with other elements in the same groups. The groups are one level higher than the individuals that are members of the groups.

We can make several points about these groups.

(1) We are using an alternative description when we refer to these groups as structural units. If we focus on the individuals making up the group, we will be describing functional phenomena: which nodes interact with which other nodes. At the group level of analysis, however, we have descriptions of structure. Thus, function at one level becomes structure at another. This type of alternative description is consistent with the logic of complex systems.

(2) This model is consistent with the one proposed by Herbert Simon (1973). He suggests that multiple-leveled hierarchical

systems, where the system is made of parts, which are made of smaller parts, and so on, are inherently more stable than monolithic non-hierarchical ones. Not only are they more stable, but they evolve faster and function more efficiently from certain points of view. One reason for these advantages can be seen in a structural aspect Simon calls loose horizontal coupling in production systems. Within a unit, there are relatively strong connections among the components. Connections between groups, however, are of a different sort. Compared to the number of intra-group links, there will be a much smaller number of inter-group links. These "bridge" links may be made by specialized linking components. At the higher level of groups, many of the microscopic details that were relevant at the individual level become irrelevant. Because of this selective loss of detail, it is possible to coordinate a number of groups into a larger system than would be possible without a hierarchical structure. To understand the function of the whole system, it thus becomes necessary to first examine its structure, as multiple-leveled structures demand similarly multiple-leveled descriptions.

In order to understand systems with multiple levels, we have to examine them at several levels of analysis. The GASSICS procedure, with its iterative looping, is set up to do just this. Although none of the single steps of the procedure include more than two levels other than the whole system, the procedure, with its repetitive loops, includes as many levels as there may be. This is how simple methods, which by themselves cannot be used in complex explanations, are put together in a procedure which theoretically has no limits in terms of multiple levels.

(3) When this model is applied to communication networks, it is consistent with the classical sociometric model that has been in use since the forties. In the sociometric model, the groups are called "cliques," and the linking components are called "liaisons" or "bridges." Traditional sociometric methods were concerned with the location of cliques and liaisons, and with the description of their function in the overall network.

(4) Finally, this model is also consistent with an information processing approach. In this approach, individual nodes would be assumed to have finite abilities to process information. As the number of nodes in a group increases, the amount of information that must be processed in the coordination of the members also increases. Clearly, this places a limit on the size of the group. If a very large number of nodes must be coordinated to perform a very complex task, a hierarchical structure will be more efficient, in terms of the amount of energy that must be expended coordinating members of the system, relative to the amount that may be devoted to achieving the goals of the system.

SUMMARY

In this chapter we developed the statistical tools needed to measure deviations from random organization. This is the second step in a cycle of the GASSICS procedure for analyzing complex systems.

This method is based on an analysis of the constraints which may be inferred to be operating in a system. We were interested here in a particular kind of constraint, one which influences patterns of interaction among the parts of the system. We showed how this kind of constraint could be operationalized by examining sets of relationships between parts in such a way as to be able to determine how much interdependence there is between all the relationships. We showed how to calculate the amount of interdependence expected in a random or unstructured system, and then how to compare the amount observed in a real system to that amount, gaining a measure of the amoung of structure beyond that expected by chance.

We observed the implications of this kind of structuring by taking a random system and rearranging the set of interrelationships between its parts in such a way as to increase the interdependence or constraint. The higher the interdependence, the more clearly organized the system became, in terms of a differentiation into intermediate level parts, which we called groups. A low-level system that is constrained or structured in this way is likely to be organized hierarchically. The system as a whole is likely to be differentiated into parts which may be further differentiated into smaller parts, and so on. This kind of arrangement is what we should look for if the system is structured. Since we can easily calculate the amount of structure, we can tell whether or not we are likely to find such a hierarchical arrangement.

This concludes the second step in a cycle of the GASSICS procedure. The next chapter will discuss the third step.

CHAPTER EIGHT

THE IDENTIFICATION OF DIFFERENTIATED SUBUNITS

Chapter Six described methods for collecting relational data about the elements in a complex system. This was the first step in a cycle of the GASSICS loop. For an applied example, we used the case of communication networks in large social systems. Chapter Seven developed the statistics needed to make the structural measurements of the second step in the GASSICS loop. In this chapter we present an algorithm for performing the third step in the loop — the identification of differentiated subunits in the system.

We begin with a general discussion of the problem and move on to a presentation of a more complete specification of the goals of this step in the analysis. Following this, we describe an algorithm that will produce the desired results. The last section presents a specific implementation of this algorithm which has been used to study structuring in human communication networks.

INTRODUCTION

In this chapter we will be discussing an algorithm for performing a certain type of analysis. The third step of the GASSICS loop is to identify differentiated subunits in the system.

Terminology

In keeping with the distinctions made in Chapter Six, we use different terms to refer to the data than the ones we use to refer to the system. The analytic terms that replace system, element, and relationship are network, node, and link. Since this step of the GASSICS loop is involved with an analysis of the network of links between the nodes, we call it network analysis. To differentiate this type of network analysis from others, and to remind ourselves that it is a part of the more general GASSICS procedure, we call it G-Network

158

Analysis. The use of this terminology will make it easier to be specific when referring to different types of analysis and to make it clear which phase of the GASSICS procedure we are referring to.

Goals of G-Network Analysis

The goals of G-Network Analysis (GNA) are fairly straightforward: to identify any existing clearly differentiated subunits in the network. These subunits will correspond to units in the system one level above the level of the parts whose interactions were described in the first step of the GASSICS loop.

Our analysis of the implications of structuring in Chapter Seven suggested that these subunits, which we called "groups," would be composed of sets of interacting low-level parts, or nodes. The nodes in any one group would be distinguishable from other nodes in the network by the fact that most of their interactions would be with other nodes in the same group. This distinction is made possible by a difference resulting from the operation of some form of organizational constraint in the original system.

1. Linkage versus distance representation

It should be emphasized here that the basis for assignment of nodes to groups is patterns of interconnections. Clearly, this conceptualization demands an analytic technique which looks at the patterns of interconnections, as actually observed, rather than some other representation. In Chapter Two we described several types of "network analysis." In many of these, data based on methods that calculate "distances" or "similarities" among the elements of the system were used. These distance methods may all be constrasted with topological or linkage-based methods such as the one demanded by the conceptualization presented here. Although some linkage-based methods, as well as many distance-based methods, require the data to be expressed in binary (either there is a relationship or there is not one) form, the method described here does not impose this constraint. On the contrary, it requires relationships to be scaled on a continuum of strengths, so that stronger relationships are represented by stronger links. The strength indicators are used in many parts of the actual analysis.

Our model is not concerned with such concepts as proximity or similarity or multidimensional distance, except as they are indirectly implied by the patterns of interconnection. Although these concepts may be used in some descriptions of the system, they are irrelevant to the model presented here. With distance models, it will be recalled, the original data are first transformed into a distance or similarity matrix, which is used as the starting point for the analysis. The "distances" (or "similarities") used in these methods are indirectly related to the actual behavior of the elements in the system. As a result, the distance matrix may be related to the raw data in very complex ways that are difficult to express in terms of actual interactive behavior. In contrast, GNA, the present method, looks only at actual patterns of interaction among the members of the system.

2. The classification of nodes

The analysis of structure in Chapter Seven suggested that we should look for clusters of elements having most of their interactions with other elements in the same clusters. The main goal will be to identify these clusters or groups.

(a) Non-participants and participants. Before we do this, however, we can eliminate all elements that do not interact at all with the rest of the system. We can also eliminate all elements that are connected in such a way that they can only function as sources or final destinations of any information or influence flowing through the system. These elements function as if they were outside the system — either parts of the environment or representatives of other systems. Although they may be important as they function in this role, they do not participate as members in the activities of the rest of the network. The elements eliminated are called non-participants.

They include both nodes not linked to the rest of the network and nodes linked only minimally. This definition will be expanded upon in a few paragraphs.

The remaining elements — the participants — will comprise the bulk of the system. These will include all the members of groups as well as the elements working to connect the groups.

The classification scheme we are using here is shown in Figure 1a. The first division is between participants and non-participants. Non-participants are subdivided into the isolates and tree nodes. There are three types of isolates — Type 1, Type 2, and Isolated Dyads.

The participants are divided into group members, liaisons, and others. The category of "linkers" includes the liaisons and others as well as any group members who have links with members of other groups. Those group-member linkers are often called "bridges."

(b) Groups. A group will have at least three members because with two it is simply a dyad, in which there is little chance for coordination with the rest of the network because then the dyad would be part of a larger group. All the members of a group must have most of their interactions within the group. There must be some path, lying entirely within the group, from each member to every other member. This insures the possibility of communication and coordination of all the members. Without this connectedness, the group cannot function as a unit and must be broken down into two or more separate parts.

We might add further restrictions to guard against situations in which a set of nodes identified as a group is really two or more groups, joined by a small number of connections, either directly or indirectly, to other groups in the set. To do this, all we have to do is stipulate that no group may be made dis-connected by the removal of some small subset of its members. If the removal of a small subset of members causes the group to become dis-connected, the group should be split apart.

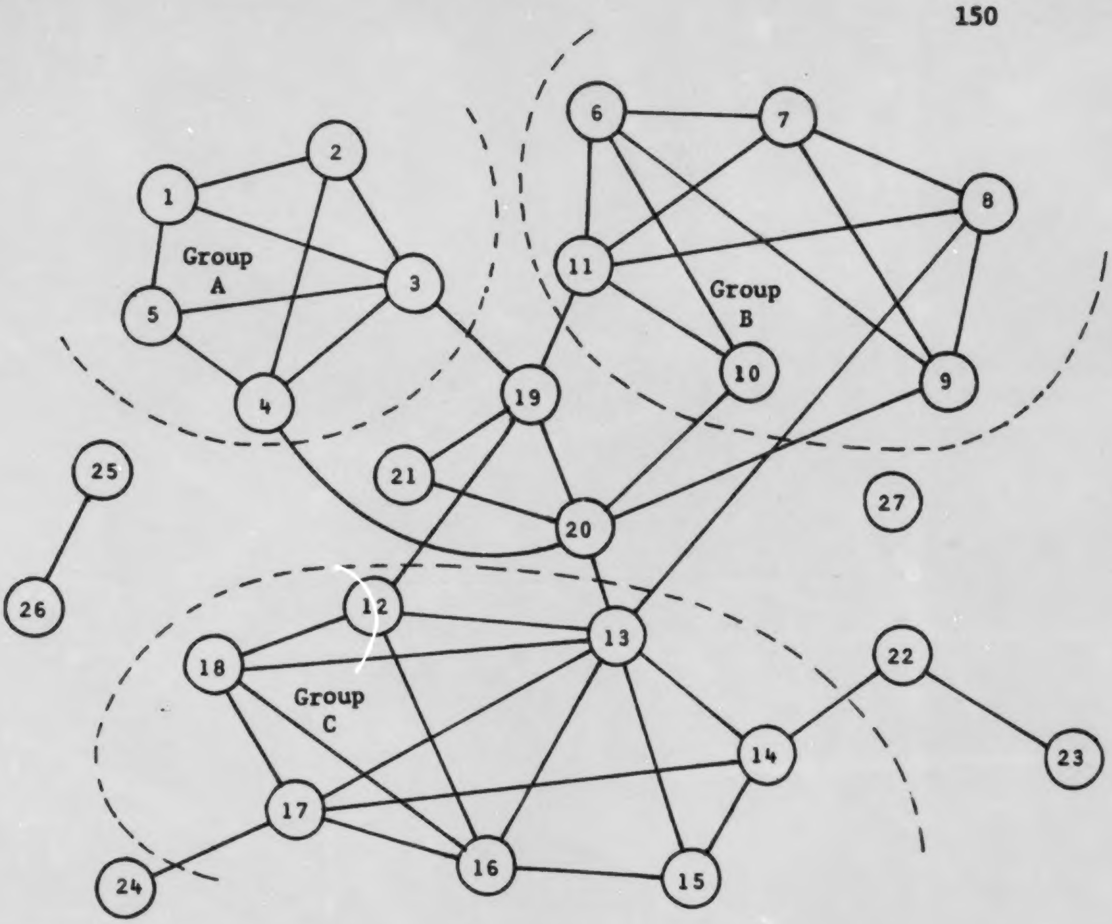

1
2
3
4
5
Group A
6
7
8
9
10
11
Group B
12
13
14
15
16
17
18
Group C
19
20
21
22
23
24
25
26
27

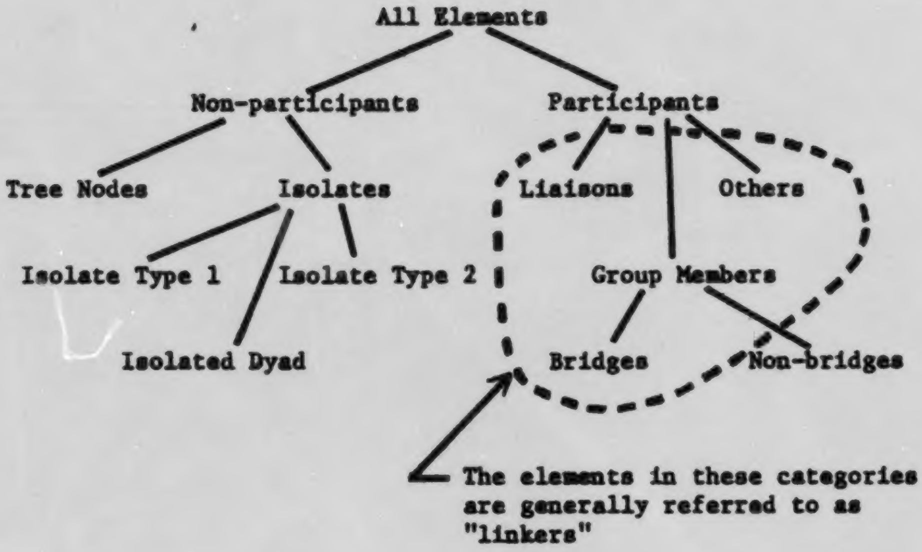
Figure 1a

NETWORK ROLES

Group Members	Linkers	Non-Participants
Group A: 1,2,3,4,5 Group B: 6,7,8,9,10,11 Group C: 12,13,14,15,16,17,18	Liaisons: 19,20 Others: 21	Isolate Type 1: 27 Isolate Type 2: 23,24 Tree Node: 22 Isolated Dyad: 25,26

Figure 1b

CLASSIFICATION TREE OF NETWORK ROLES


All Elements
Non-participants
Participants
Tree Nodes
Isolates
Liaisons
Others
Isolate Type 1
Isolate Type 2
Group Members
Isolated Dyad
Bridges
Non-bridges
The elements in these categories are generally referred to as "linkers"

Thus, the criteria that must be satisfied in order for a set of nodes to be called a group are:

(i) There must be at least three members.

(ii) Each must have most of its linkage with other members of the same group. Here, amount of linkage refers to amount of interaction rather than number of relationships. This point is clarified in a few paragraphs.

(iii) There must be some path, lying entirely within the group, from each member to each other member (this is called the connectedness criterion).

(iv) There must be no subset of links or nodes which, if removed, causes the group to become dis-connected (in practice, the subset is usually smaller than ten percent of the group — this is called the critical links/nodes criterion).* The nodes that are removed when testing this criterion are examined to see which role they fit (they typically become liaisons or group members).

(c) Participants -- group members and linkers. Participants are then either group members or "linkers," connecting the groups. These role categories are defined by the following criteria.

(i) group member -- a node with most of its linkage with other members of the same group.

(ii) liaison -- nodes which have most of their linkage with group members in general, but not with members of any single group.

(iii) type other — nodes which do not have most of their linkage with group members in general.

Liaisons and others are the "linkers" mentioned above. The difference between the two is that liaisons work directly with the groups they connect while others do not. Others can be thought of as providing more indirect, or multi-step, links between groups. A third type of linker is a group member which has connections with members of other groups. These nodes are called "bridges." Non-bridge group members, bridge group members, liaisons, and others may be compared on their patterns of connections with groups' members:

---

*This criterion has been stated as two separate criteria in the past (Richards, 1974); one for critical links and one for critical nodes. Operationally, the same procedures are done when testing the separate criteria, so they have been stated as a single criterion here.

> Non-bridge group members have most of their linkage with members of their own groups and no linkage with members of any other group.
>
> Bridge group members have most of their linkage with members of their own groups and may have some linkage with members of other groups.
>
> Liaisons have most of their linkage with members of groups but not with the members of any single group.
>
> Others have less than half of their linkage with members of groups. (Most of their linkage must be with liaisons and other others.)

(d) Non-participants. We can provide similar criteria for the classification of non-participants.

(i) Isolate Type 1, or Unattached Isolate — these nodes have no links whatsoever.

(ii) Isolate Type 2, or Attached Isolate — these nodes have only a single link and thus cannot take part in the transfer of information or influence through the network. They may, however, function as sources of information if they either are outside the system or have links outside the system.

(iii) Isolated Dyad -- these nodes are similar to attached isolate pairs who are linked to each other. In terms of contact with the rest of the network, they function more like Isolate Type Ones.

(iv) Tree Node -- sometimes there is a chain of isolates where each one is attached to the next and only one end is connected to the rest of the network by a contact to a participant. The node at the other end of the chain will have only one link and will be an attached isolate. The nodes in the chain between the attached isolate and the participant are all called

> Tree Nodes. If any link in the tree structure is cut, the distal part of the cut (the part farthest from the participant end) will be separated from the rest of the system. All the categories of nodes are illustrated in Figure 1b.

(e) Amount of linkage versus number of links. Several of the criteria refer to a proportion of a node's linkage. We are referring here to the amount of linkage, rather than the number of links. The amount is operationalized in terms related to interactions among the parts, where interactions are viewed as processes involving constraint. We are interested here in the amount of constraint rather than the number of constraining relationships. To give a specific example: if the amount is operationalized as "time spent interacting" when the network is for a communication system, we would look at the appropriate set of links in terms of what fraction of the total amount of time they comprise, rather than in terms of how many links there are. Thus, a node having ten links could be a group member even if only two of those ten links were with members of the group, as long as those two links account for more than half of the total linkage. For example, those two links might take four hours per week. If the other eight links combined total less than four hours per week, the node would be a member of the group.

The goals of GNA are thus to classify the nodes in the network into the various network roles, based on their patterns of interaction with each other and to provide as much information as possible about the system at each of three levels — the individual node, the group, and the entire set of nodes.

Taken by itself, out of the context of the GASSICS procedure, GNA operates only at three levels. In the context of the larger framework of GASSICS, however, it may be used to provide information at several additional levels. The issue is discussed more completely in the next chapter.

AN ALGORITHM FOR PERFORMING G-NETWORK ANALYSIS

So far we have seen what the goals of GNA are and what kind of data it uses. In this section we will see how the actual analysis is accomplished. The main tool is NEGOPY, the GNA program. In the first part of this section we will discuss the algorithm upon which the program is based because an understanding of the procedures carried out by the program is necessary to any potential user of the program, in order to know what is happening to the data. In Part I of the Appendix we will take up several considerations directly related to the 1975 CDC implementation of the program, such as limits on the data, specific input requirements, and so on. We will also discuss the various options the user has when running NEGOPY and the output of the program -- what the various tables mean and how to interpret them. Detailed information on the actual use of the program -- how to prepare control cards, error messages, and how to interpret them, how to "fine tune" the program, and several other miscellaneous issues -- are discussed in Part II.

The point was made earlier that GNA is a topological method -- it looks for specific patterns in the data. The realization that this is a pattern recognition problem made it possible to program a computer to do the analysis. The GNA program, then, is based on a pattern recognition algorithm. Although it uses a variety of statistical and mathematical operations as it carries out an analysis, it is not based on mathematical or statistical procedures, as are other kinds of analytic programs, like, for instance, factor analysis.

In addition to this difference, the GNA program works with information from not one but three levels of analysis at one time. In GNA there are not only individual nodes but also groups of nodes and the whole system. The differences between this kind of network analysis and other types of analysis will be discussed in depth in Chapter Nine. We move now to a discussion of the algorithm upon which the GNA program is based.

There are five steps in the analysis. First, the data are read in, cleaned, and organized in an orderly fashion. Second, an iterative operation that makes the actual pattern recognition part possible is performed. In the third stage, the pattern recognition algorithm is carried out. Here, groups are tentatively identified. In the fourth stage, the strict criteria for the various role definitions are applied and the tentative solution produced earlier is tested and made exact. The results of the analysis are printed out in the form of various tables and charts in the fifth and final stage. In the following section, we will be concerned only with the actual computational parts of the analysis. Thus, we will cover only the second, third, and fourth stages here. The other stages are covered in Part I, where NEGOPY, the actual computer program, is described.

The Classification of Non-participants

Before beginning the group detection routines, all non-participants are identified and classified in the appropriate way. Unattached isolates (Isolates Type 1) are located first. These are the nodes with no links whatsoever. Isolated dyads and attached isolates (Isolates Type 2) are identified next. These nodes all have exactly one link.

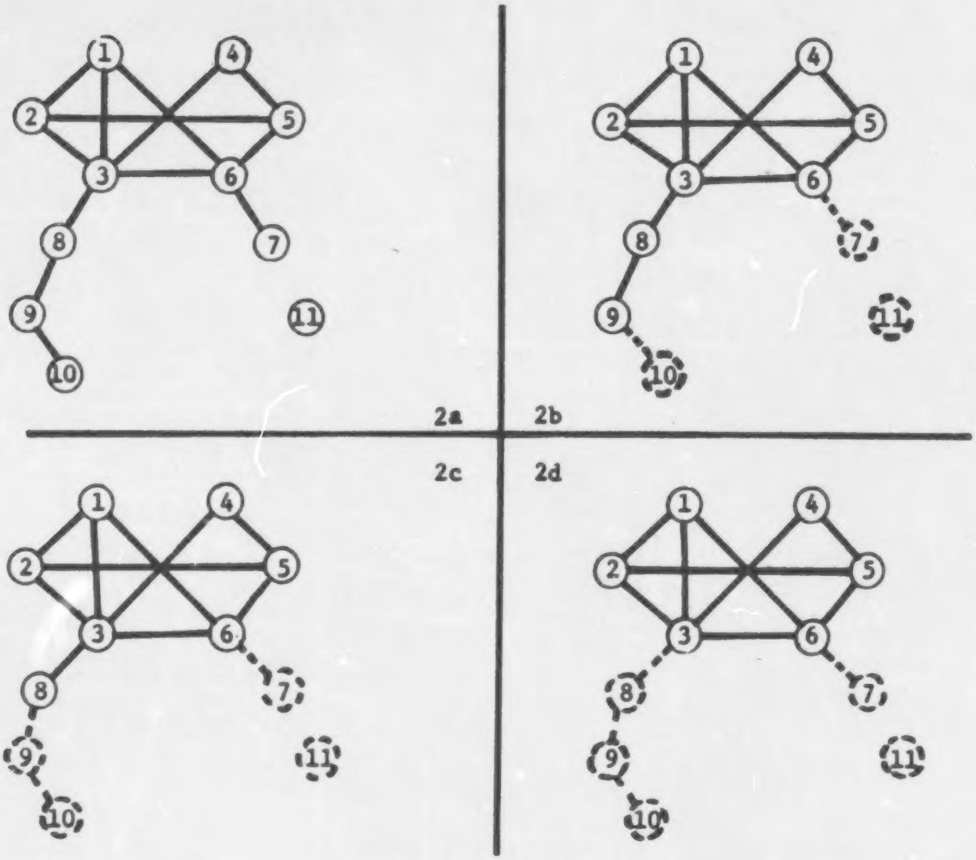
The next step is to identify the tree nodes. If links to attached isolates are ignored, some tree nodes will be left with only one link. These are one-step tree nodes. If links with one-step tree nodes are also ignored, two-step tree nodes will be the only ones having one link. This process continues until all tree nodes are identified. The process is shown in Figure 2.

The Algorithm Which Identifies Groups

The major task to be accomplished in this part of the analysis is to identify the groups. We have data describing the relationships between the individual nodes. If we can represent the data in the right way, it will be easy to "see" the groups. The representation

Figure 2

THE IDENTIFICATION OF NON-PARTICIPANTS



2a shows a part of a network. Node 11 is an unattached isolate; 10 and 7 are attached isolates; 8 and 9 are tree nodes, and the others are group members. None of the nodes have been classified in 2a. In 2b node 11, the unattached isolate, has been identified by the absence of any links. 7 and 10, the attached isolates, have also been identified by the fact that each has only one link. In 2c, node 9 has been identified as a one-step tree node. This identification was made because the node has only one link, not counting links to isolates. In 2d, node 8 has been identified as a two-step tree node. 8 has only one link, not counting links to already identified non-participants.

we would like to have would be one in which the members of each group as "close" to other members of the same group and "far" from the members of other groups. Then we would just "look" for clusters — groups -- sets of nodes having most of their linkage to other nodes in the same groups. This will be a graphical representation of the data -- nodes will be "moved around" until their locations, relative to other nodes, can be used to decide the way they fit into the network.

1. The vector averaging process

The first step in the identification of groups is to rearrange the data so that the groups become visible; the second is to identify the groups. The way the first step is accomplished can be understood with the following analogy. Imagine the nodes to be like billiard balls scattered about in space. Imagine there to be rubber bands connecting the balls corresponding to nodes with links between them. Imagine there to be springs between balls corresponding to nodes that do not have links between them. The rubber bands will act to pull the balls connected to each other closer to each other, while the springs will push the balls not connected to each other apart from each other. If we hook up the rubber bands and springs and release the balls, they will rearrange themselves so that the balls corresponding to nodes with links to each other will be close to each other, while the balls corresponding to nodes that are not linked to each other will be pushed away from each other. This example is shown in Figure 3.

We could refine this technique by using heavier rubber bands to represent the links that occur more often or that are more important. Since our objective here is to make it easier to identify groups, we could make the process operate even better if we could make the rubber bands for within-group links heavier than the ones for other kinds of links. In order to do this, we need some indicator that tells us which links look like within-group links.

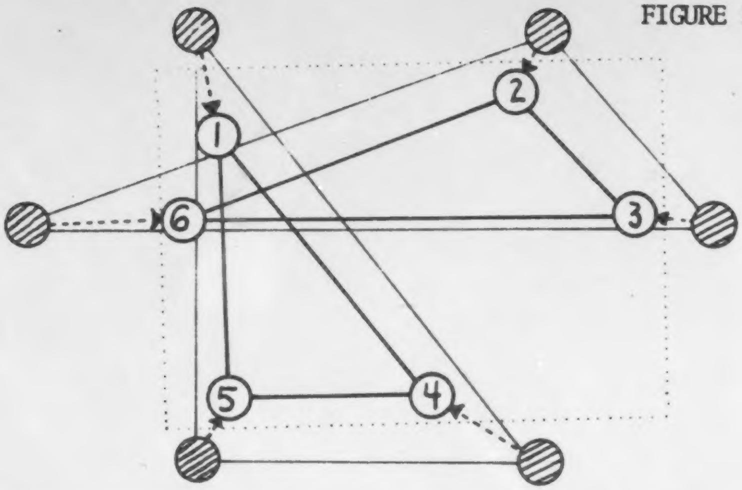
Figure 3

THE "BALLS AND BANDS" MODEL

This figure illustrates the billiard ball and rubber band model described in the test. The network shown has two groups of three nodes each. The three drawings represent three successive increments of time, as the nodes move farther and farther in response to the forces exerted by the rubber bands.

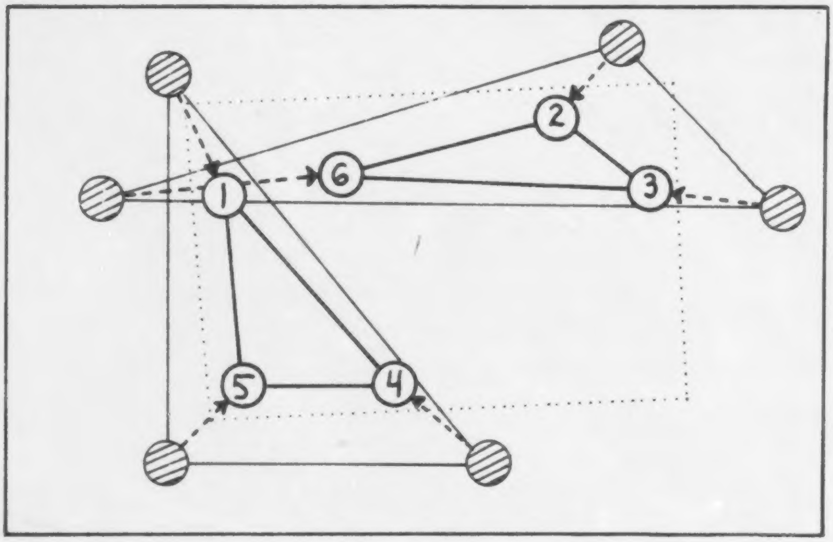
The original position of the balls is shown by the shaded circles in the top drawing. Movement of balls during each time increment is shown by the dotted arrows in the three drawings. The scale was changed in going from the first to the second to the third drawing in order to show smaller and smaller regions in space as occupying the same sized area in the drawings. The region of the top drawing shown in the middle one is indicated by the dotted box in the top. Similarly, the area of the bottom drawing is shown by the dotted box in the middle one.

FIGURE 3



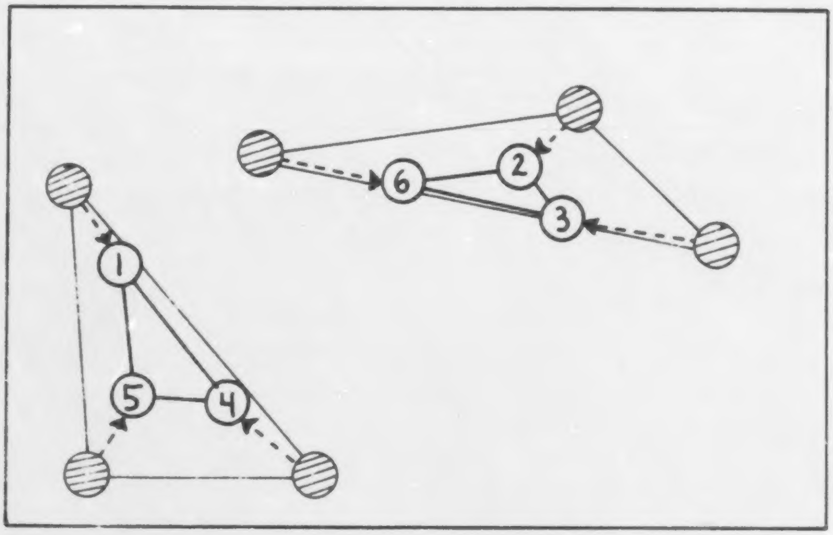
SHADED CIRCLES - NODES AT TIME 0

NUMBERED CIRCLES - NODES AT TIME 1



SHADED CIRCLES - TIME 1

NUMBERED CIRCLES - TIME 2



SHADED CIRCLES - TIME 2

NUMBERED CIRCLES - TIME 3

If two nodes are in the same group, they are likely to have many links to the same nodes. There is likely to be a high number of shared links, or two-step links, between this pair of nodes. If they are not in the same group, they are not likely to interact with the same nodes and there are therefore not likely to be many two-step links between the nodes. Thus, the number of two-step links is used by the program as an indicator of the probability that the link is a within-group link.

Now, it is difficult to represent large numbers of points in multi-dimensional space. It takes a lot of information to do so, and it is fairly difficult to move objects in this kind of space. Extensive experimentation with real data, however, showed that it was not necessary to use a multi-dimensional representation for this analysis; a single line segment was sufficient. This kind of reduction in complexity of representation both reduced the amount of information needed to perform the analysis and made the analysis itself become easier to do.

The analysis is performed as follows: nodes are scattered at unit points along a line segment N units long, where N is the number of nodes. We then treat each link from, say, node A to node B as a vector, starting at A and pointing at B. We take all the vectors for each node and compute the average, weighting the individual vectors for strength of the link and probability that the link is a within-group link. We then get a single point for each individual node, that point being the mean of that node's vectors. This is illustrated in Figure 4. After all the means have been computed, each node is moved to the point indicated by its mean.

After this process has been completed, nodes with links to each other will be closer to each other than they were before. They will not, however, be as close as they could be. This fact is due to the way nodes are scattered initially, and also because of the statistical properties of the mean. For this reason, the entire process is repeated, using the new locations instead of the original positions used for the first set of calculations. A plot showing how the nodes move in successive iterations is shown in Figure 5. Between each set

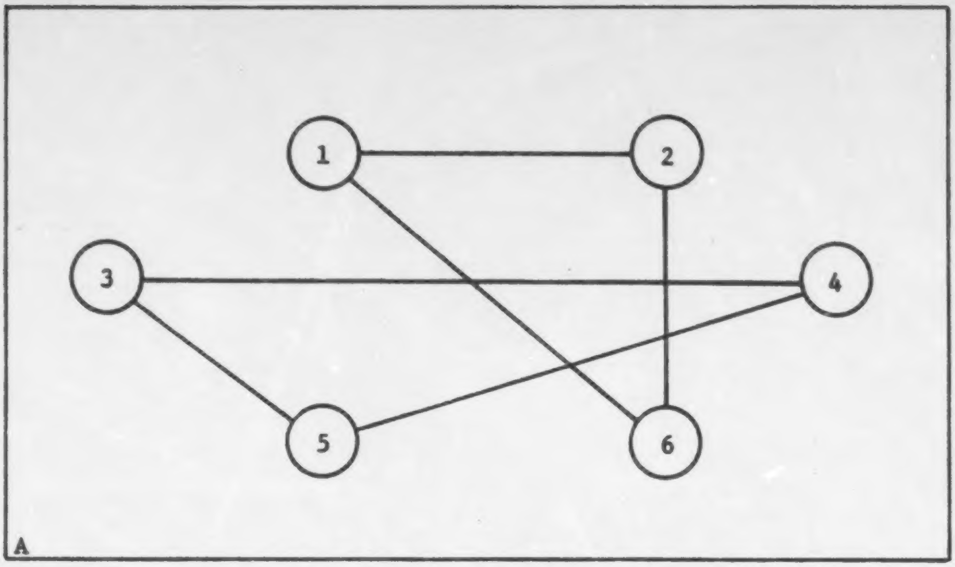


Figure 4

THE VECTOR AVERAGING PROCESS

In A above is shown a hypothetical network, consisting of two groups, each of which has three members. The vector averaging process is illustrated in B on the next page.

The first set of three "arrows" shows the process for Node #1. On the top is the vector for the link from #1 to #2. Below that is the vector for the link to #6.

The solid circle on the third line is the starting point for Node #1. The mean for #1 is calculated by adding together the two vectors -- to #2 and #6. The location of Node #1 is also added. The sum is then 2 + 6 + 1 = 9. The mean is simply 9/3 = 3. Node #1 is moved to the mean, the point shown as an open circle on the third line.

The next five sets of lines show the process for nodes 2, 3, 4, 5, and 6, respectively.

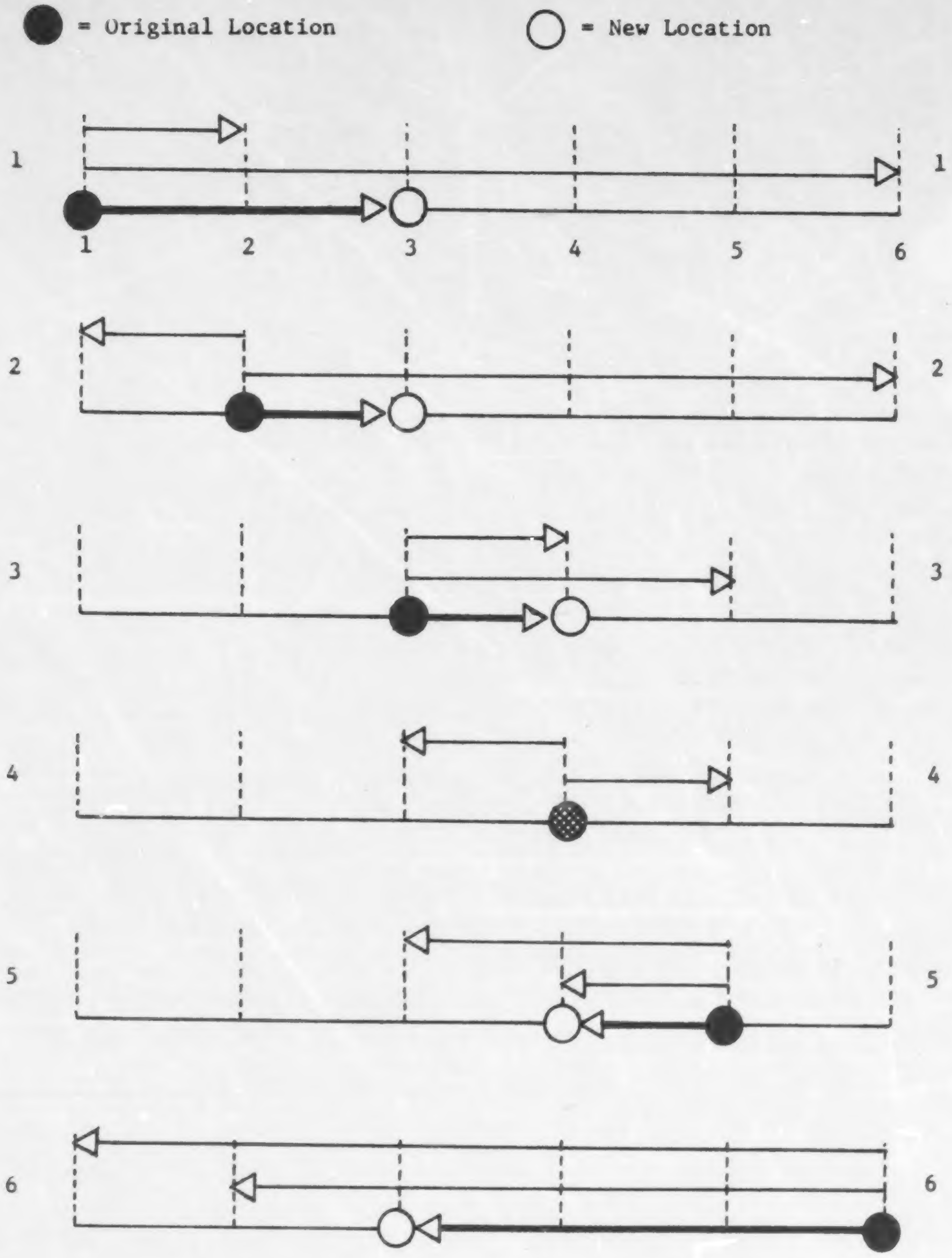
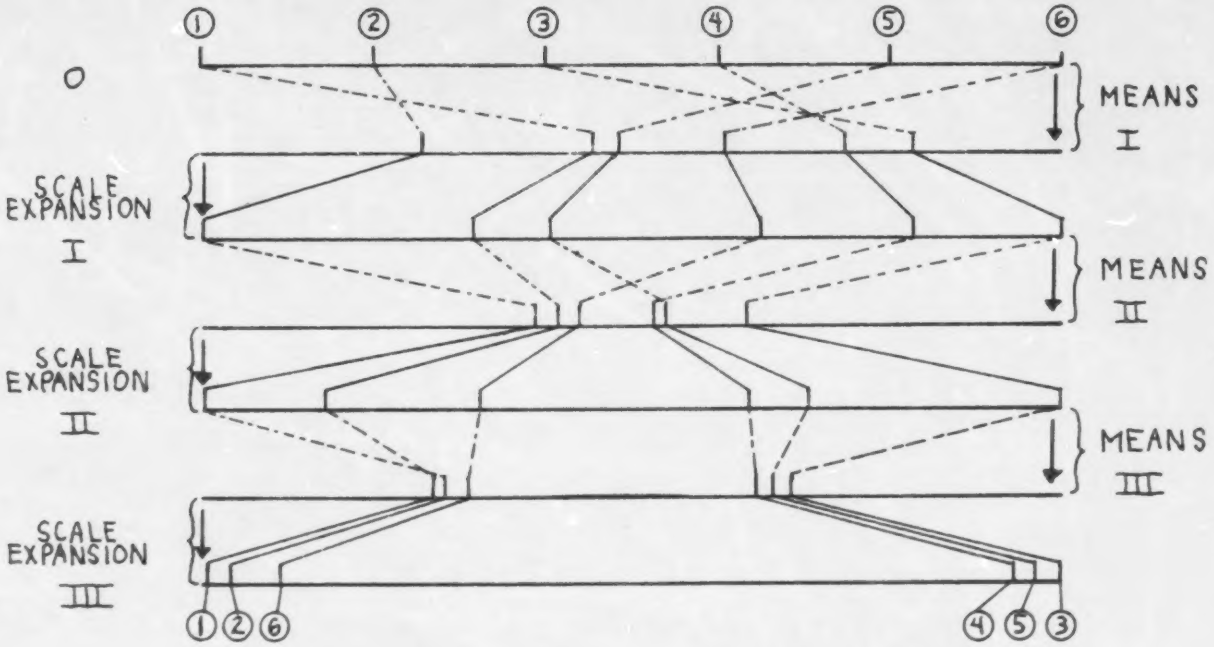

● = Original Location
○ = New Location
1
1
1
2
3
4
5
6
2
2
3
3
4
4
5
5
6
6

Figure 4b

Figure 5

THE ITERATIVE PROCESS OF VECTOR AVERAGING


①
②
③
④
⑤
⑥
O
MEANS I
SCALE EXPANSION I
MEANS II
SCALE EXPANSION II
MEANS III
SCALE EXPANSION III
①
②
⑥
④
⑤
③

This diagram shows how the iterative process of vector averaging works. The first line shows the initial positions of the six nodes. The second shows what the means could look like. Moving from the second to the third lines, the scale has been expanded so that the nodes range over the entire length of the continuum. The fourth and sixth lines show the second and third sets of means, while the expanded versions are shown on the fifth and seventh lines. (Note that the values shown are not the actual values that would be obtained for this particular network; they are intended merely to illustrate how the process might typically look.)

of calculations it is necessary to expand the scale of the continuum so that the spread or range which is occupied by the nodes remains N units long. If this is not done, the points will move closer and closer to each other, finally collapsing on a single spot. This is the "scale expansion" referred to in Figure 5.

The formula used for calculating a node's mean is shown here:

$$M' = \frac{\Sigma(wf_i \cdot S_i \cdot M_i)}{\Sigma(wf_i \cdot S_i)} ,$$

where wf_i is the two-step weighting factor described above (the number of two-step links connecting the nodes plus one); S_i is a ratio-level indicator of the strength of the link; and M_i is the old mean of the node to whom the link goes. The summation is done as i goes from 1 to ℓ, where ℓ is the number of links that the individual node whose mean we are calculating has.

In the development of this algorithm, different numbers of iterations, different ways of varying relative contributions of wf_i's, S_i's, and M_i's, and different ways of assigning the original M_i's were tried. In general, four to six iterations seemed to be sufficient for any data set that was examined. If nodes are given "subject numbers" running from 1 to N, where N is the number of nodes, and these subject numbers are used as the first approximation for the M_i's, the process seems to work well for all types of data. In actual tests, when different subject numbers were assigned to nodes, the solution obtained was identical to the first solution, which indicates that the process is not terribly sensitive to the original positions. Usually, the wf_i's and S_i's are given equal weight, although this has not been tested extensively. The computer program allows the effects of the wf_i's to be varied by the use of an additional weighting factor W2S, which is shown in the equation below. When it is set to one, both the wf_i's and the S_i's have equal influence.

$$M' = \frac{\Sigma((W2S \cdot wf_i) \cdot S_i \cdot M_i)}{\Sigma((W2S \cdot wf_i) \cdot S_i)}$$

The result of the application of this process is a continuum, N units long, with a scattering of nodes along its length. Because of the way it was obtained, the continuum cannot be directly interpreted in terms of the behavior of the individual nodes in the network. It is impossible to determine the relation between specific raw data and the final results for any given node in complex iterative procedures like this one. The relationship is statistical and must be interpreted in statistical terms: the nodes that are close to each other on the continuum are likely to be somehow "close" in the network. That is, they probably are in the same group. It is impossible to say anything specific about whether or not the nodes are linked or how many two-step links there are connecting the nodes. A sample network, together with the continuum that might result, is shown in Figure 6. This continuum is used as the input to the next stage of the analysis, in which tentative boundaries for groups are drawn.

2. Drawing the tentative boundaries

For any human observer, even a casual glance at Figure 4 will be enough to suggest that there are three clusters of nodes. The computer, however, must be "told" what a cluster looks like and how to "look" for one. People probably identify a cluster as an area in which there are many nodes surrounded by areas in which there are fewer nodes. This is essentially what we have the computer "look for."

We will need a plot of the "density" of nodes along the continuum. In order to obtain such a plot, we construct a "window" and move it along the continuum, counting the number of nodes visible through the window at each point. This is shown at the top of Figure 5. The optimum size of the window, determined by experimentation, appears to be about two units on an N unit line. Windows smaller than this introduce spurious statistical information, while with windows larger than this group boundaries tend to blur and merge into indistinction. This is shown in Figure 5, where density plots appear for windows of varying widths. The result of moving the window down the

Figure 6

A NETWORK AND THE CONTINUUM FOR THE NETWORK

The top of this figure shows a hypothetical network composed of twenty nodes. Group boundaries are indicated by the dashed lines.

The bottom shows what the final continuum might look like for the network shown in the top. Again, the group boundaries have been indicated by dashed lines.

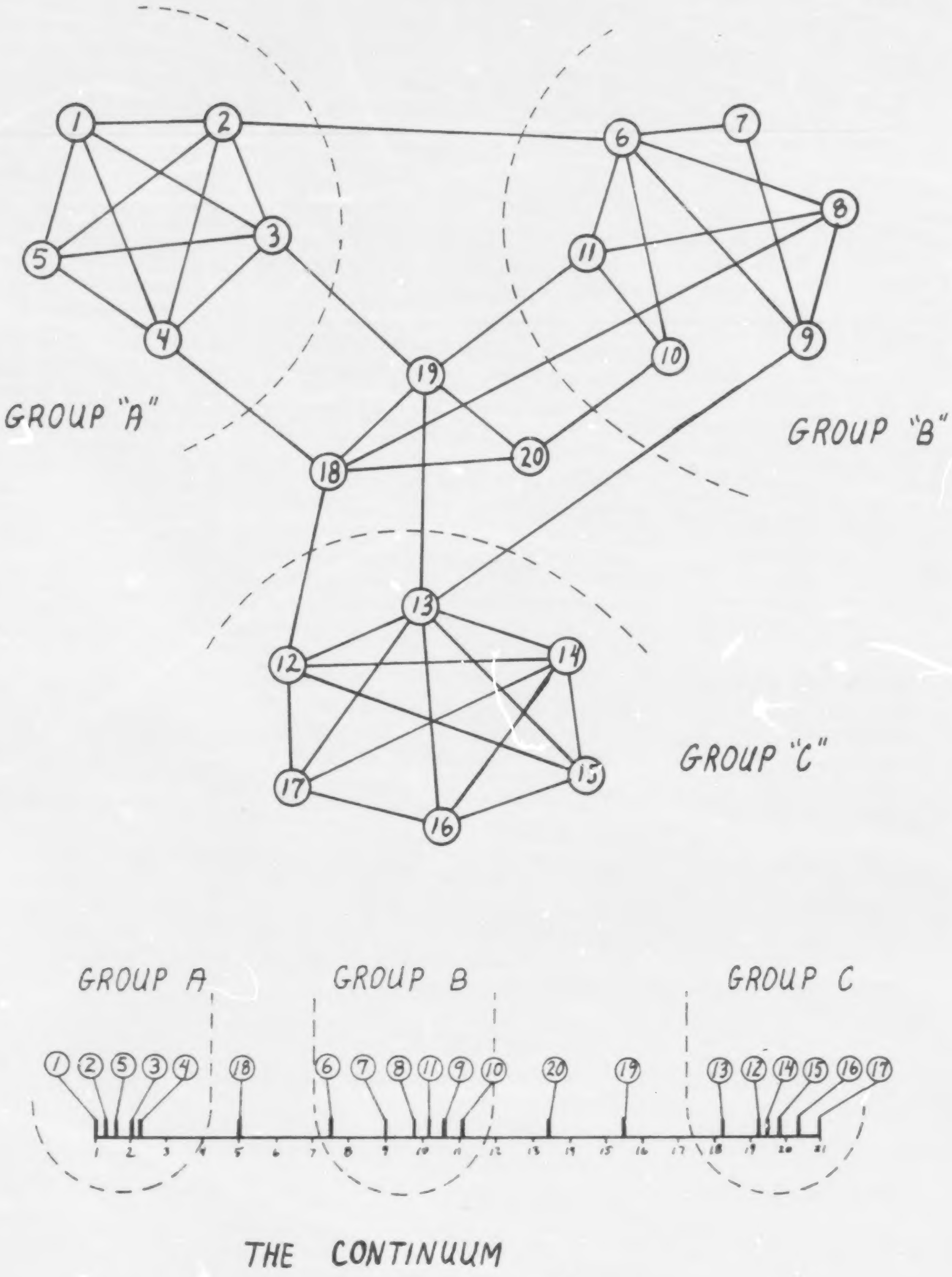

GROUP "A"
GROUP "B"
GROUP "C"
GROUP A
GROUP B
GROUP C
1 2 5 3 4 18 6 7 8 11 9 10 20 19 13 12 14 15 16 17
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
THE CONTINUUM

FIGURE 6

continuum will be a list of densities, with one value for each individual node. Such a list could be represented as a bar graph like the one shown in Figure 7.

With this representation, groups will look like mounds, with boundaries between groups being indicated by low points. Although it seems as though this representation would be adequate, there arose problems which lead to an improvement over this simple plot. Although the problems will not be discussed here, the improvement will: instead of just counting the number of nodes visible through the window, two numbers are counted -- the number visible on the right half of the window and the number visible on the left half. When constructing the bar graph, the number visible on the right half is plotted above the horizontal while the number visible on the left half is plotted below the horizontal. The result is shown at the bottom of Figure 7.

The final step in this stage is to have the computer "draw" lines around the groups. This is done by locating spots at which there is a large change in density as we move from one point on the continuum to the next. If we count the number of non-overlapping points and divide by the number of overlapping points for each pair of adjacent nodes on the final bar plot, we will have a fairly sensitive indicator of group continuity. This is shown in Figure 8. High values for this ratio will indicate that there is a large change as we move from one node to the next. Low values, on the other hand, will indicate that there is only a small change. If we choose a cutting point and instruct the computer to "draw" a line whenever the ratio goes above the cutting point, we will have "told" the computer how to "draw" the boundaries around groups. If the value of the cutting point is variable, we can alter the sensitivity of the group spotting routine in either direction. With a window of two units, a cutting point of 1.0 appears to be optimum for most networks. Different values, along with the results, are shown in Figure 8.

There is one final aspect of the process of scanning the continuum that is important here: "scan mode." The computer begins at one end of the continuum and works through the locations of every node

Figure 7

SHOWING THE OPERATION OF THE WINDOW AND THE CONSTRUCTURING OF THE DENSITY PLOT

This figure shows how the density plot is made. The example uses the continuum shown in Figure 6. In the top part, the window is shown, centered successively on the first eight nodes.

The three bar graphs in the middle show the effects of differently sized windows.

On the bottom is shown the refined version of the plot, with numbers of nodes visible to the right of the center of the window plotted above the horizontal and numbers visible on the left of the window plotted below the horizontal.

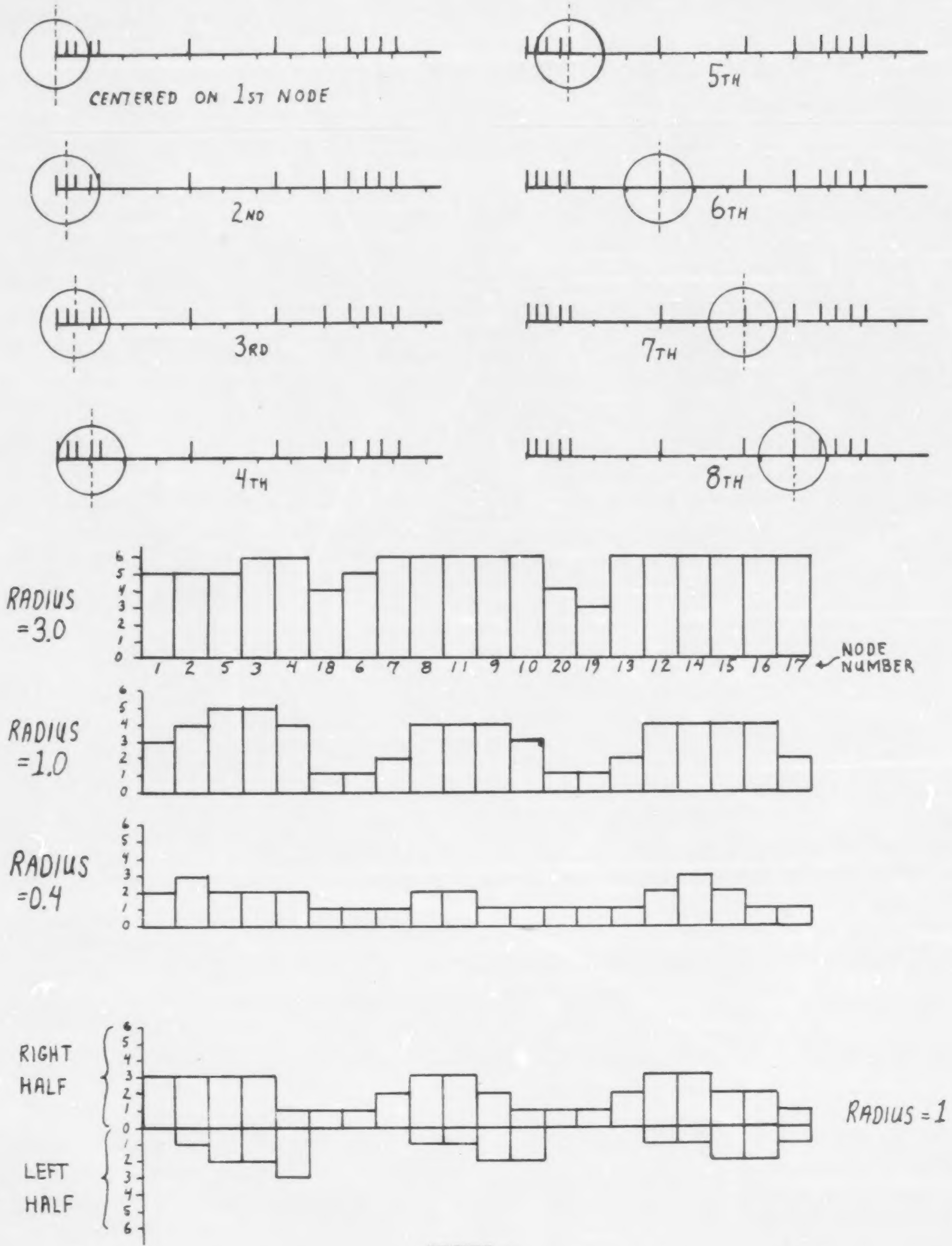

CENTERED ON 1ST NODE
2ND
3RD
4TH
5TH
6TH
7TH
8TH
RADIUS =3.0
0 1 2 3 4 5 6
1 2 5 3 4 18 6 7 8 11 9 10 20 19 13 12 14 15 16 17
NODE NUMBER
RADIUS =1.0
0 1 2 3 4 5 6
RADIUS =0.4
0 1 2 3 4 5 6
RIGHT HALF
LEFT HALF
6 5 4 3 2 1 0 1 2 3 4 5 6
RADIUS = 1

FIGURE 7

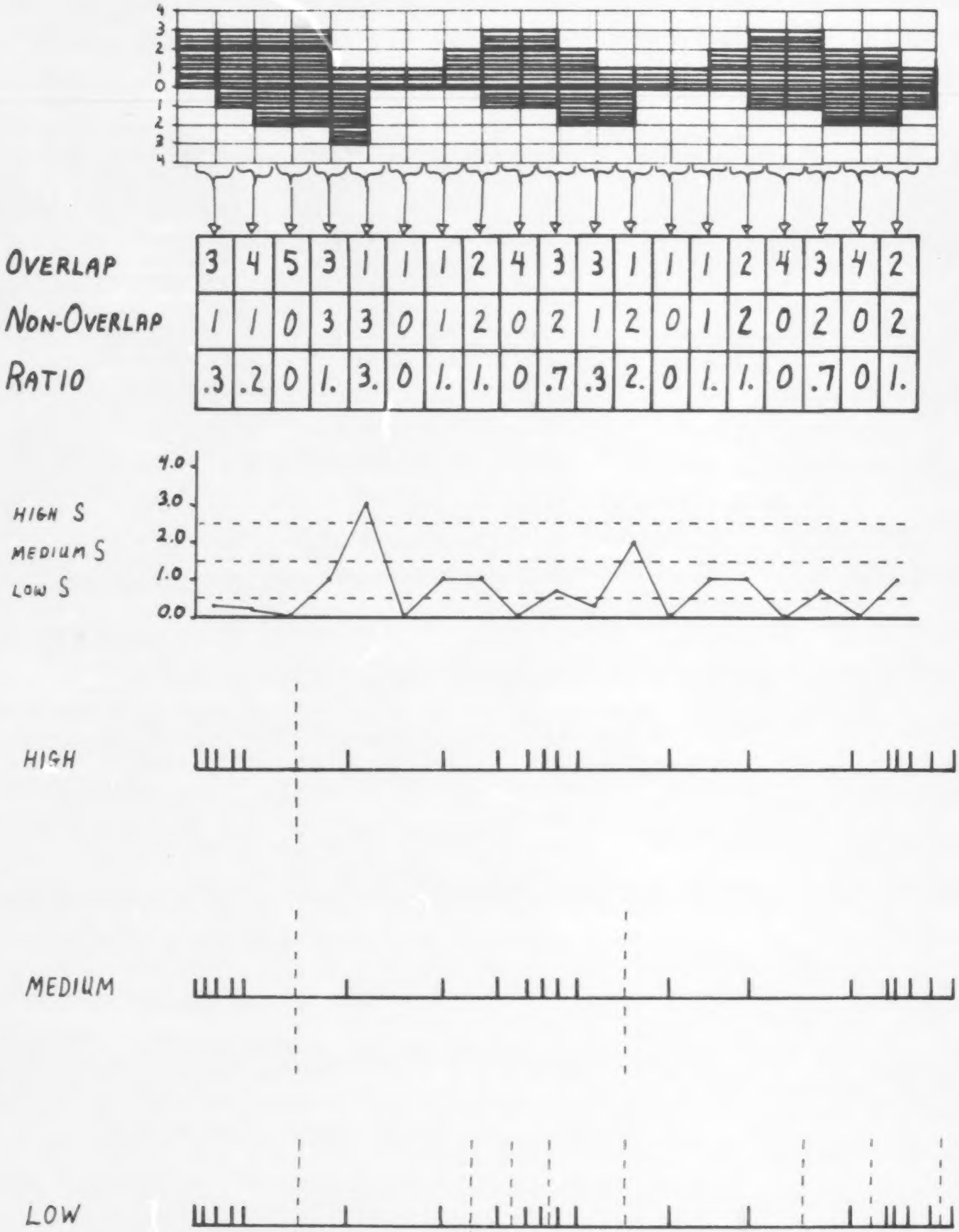
Figure 8

CUTTING POINTS AND THE BOUNDARY-DRAWING PROCESS

This figure illustrates the boundary-drawing process. The density plot on the bottom of Figure 7 is shown on the top of this figure. The table below tge plot shows the number of overlapping points, the number of non-overlapping points, and the ratio of the two numbers for each successive pair of bars on the bar plot.

The ratios are plotted in the graph in the middle of the page. The three dotted lines show the three different cutting points.

Below the ratio plot, the original continuum is shown three times. The first shows the effect of a high cutting point, while the second and third ones show the results for moderate and low values of the cutting point.


4
3
2
1
0
1
2
3
4
4.0
3.0
2.0
1.0
0.0
HIGH S
MEDIUM S
LOW S
HIGH
MEDIUM
LOW

OVERLAP	3	4	5	3	1	1	1	2	4	3	3	1	1	1	2	4	3	4	2
NON-OVERLAP	1	1	0	3	3	0	1	2	0	2	1	2	0	1	2	0	2	0	2
RATIO	.3	.2	0	1.	3.	0	1.	1.	0	.7	.3	2.	0	1.	1.	0	.7	0	1.

FIGURE 8

along the continuum. There are two "modes" the computer can be in at any point of the scan. One is "group mode" and the other is "non-group mode." When the nodes between a set of the lines drawn by the computer satisfy a set of criteria, the computer is in "group mode." This will generally be the case when the densities in the region of the nodes are high. When the density is low, the computer scans in "non-group mode." Usually, the mode will be non-group between groups — areas along the continuum in which the density of nodes is very low. Although the computer can only switch modes when lines are drawn, it does not necessarily switch every time a line is drawn. The nodes located between a pair of lines in which the scan mode was "group" are tentatively identified as a group. The other nodes are classified later on the basis of their interaction patterns with other nodes.

This concludes the approximate phase of the analysis. The result of this stage is a list of tentative groups of nodes. The next part of the analysis involves the testing of this tentative solution and any alteration that may have to be done to "clean it up."

3. Using the criteria for an exact solution

This part of the analysis can be divided into two parts. In the first, individual nodes are tested to see if they meet the relevant criteria for their role in the network. If they do not, the appropriate changes are made. In the second, whole groups are tested for the criteria that are relevant at that level. Again, appropriate changes are made if necessary. We begin with the individual testing, which is very simple.

(a) Individual testing. First, nodes not in groups are tested to see if they meet the criterion for either liaison or group membership in any group. If any individual node does meet the criterion, it is reclassified on that basis. If the node fails both tests, it is labelled as "type other."

Second, members of groups are tested to see if they meet the criterion for group membership. Again, if the criterion is not met the appropriate changes are made.

Because changes made at any point in time can affect the roles of other nodes who were tested earlier, the tests are applied twice to make sure that the final classification will be consistent with itself.

(b) Group testing. In this section, we change our level of analysis to whole groups, rather than separate nodes. The criteria to be tested in this part are the connectiveness and critical link/node criteria. Since the information generated in the testing of the connectiveness criterion is necessary in the testing of the other one, it will be covered first.

(i) Forming the distance matrix -- The basic device used in the testing of these criteria is the distance matrix, which is constructed for each group. In this n-by-n matrix (n is the number of members in the group), the entry in row i, column j gives the number of steps needed to get from node i to node j in the group. If there is some finite number in each element of the matrix, the group will be connected. This means that there will be some path from each node in the group to every other node in the group. The longest any path could be is n-1 steps. A sample network, together with its distance matrix, is shown in Figure 9.

The distance matrix is constructed as follows. A matrix is constructed in which there is a row and a column for each node in the group. All the elements are initialized to zero. Whenever there is a link from node i to node j, a "1" is entered in row i, column j. If the link is reciprocated, a "1" is also entered in row j, column i.

A Boolean logic operation which is analgoous to raising the matrix to successively higher and higher powers is then

Figure 9

A DISTANCE MATRIX

At the top of Figure 9 is shown a hypothetical eight-node network. The matrix directly below the network is a binary version of the network. In this matrix, each node has a row and a column. The i,j entry of the matrix is 1 if node i is linked to node j.

The second matrix is the distance matrix for the same network. The entry in the i,j element of the matrix is the number of links in the shortest path from node i to node j.

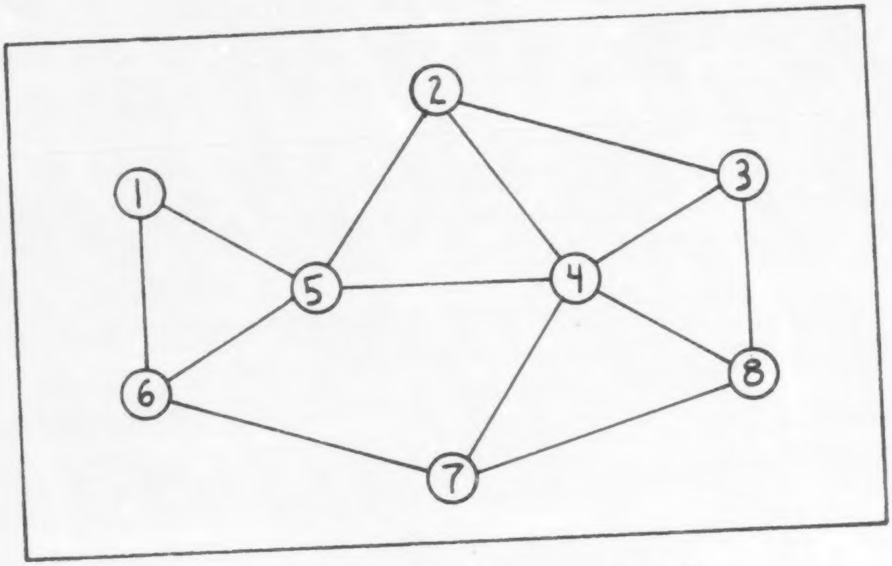


FIGURE 9

BINARY FORM

	1	2	3	4	5	6	7	8
1	0	0	0	0	1	1	0	0
2	0	0	1	1	1	0	0	0
3	0	1	0	1	0	0	0	1
4	0	1	1	0	1	0	1	1
5	1	1	0	1	0	1	0	0
6	1	0	0	0	1	0	1	0
7	0	0	0	1	0	1	0	1
8	0	0	1	1	0	0	1	0

DISTANCE FORM

	1	2	3	4	5	6	7	8
1	0	2	3	2	1	1	2	3
2	2	0	1	1	1	2	2	2
3	3	1	0	1	2	3	2	1
4	2	1	1	0	1	2	1	1
5	1	1	2	1	0	1	2	2
6	1	2	3	2	1	0	1	2
7	2	2	2	1	2	1	0	1
8	3	2	1	1	2	2	1	0

performed. Instead of setting the i,j entry in the product matrix to the value of the cross product of the ith row and the jth column, however, the first power on which this value becomes non-zero is used.

The raising of the matrix to higher powers is stopped when one of two conditions obtains either: (a) all off-diagonal elements become non-zero, which implies the group is connected; or (b) when going from any power k to the next power k+1, no entries change value, which implies the group is not connected at level k and will _never_ be connected at _any_ level.

If the group is not connected, it is split into a connected part and all the rest. Each of the two parts is then treated as a separate group, and subjected to all the tests that any group must undergo.

(ii) _Testing for critical nodes._ At this point there is only the critical links/nodes criterion remaining to be tested. This criterion serves as a check against situations like those shown in the bottom half of Figure 10, where two groups have been mistakenly identified as one. This situation is generalized to include situations in which there are any number of multiple groups, connected in some relatively minimal way, which we wish to separate into distinct groups. The occurrence of these confusions is a result of the inelegance of the approximate techniques used in the first half of the analysis. For analytic purposes, it was practical to combine both critical links and nodes into a single rule which says that no subset of some arbitrary size may be removed from a group and cause the group to become disconnected. If there is such a subset, the group will be seen to be "really" two or more groups. As a result of this combination, whenever two groups are joined by a bridge link (a link between members of different groups), one of the nodes of this link will be

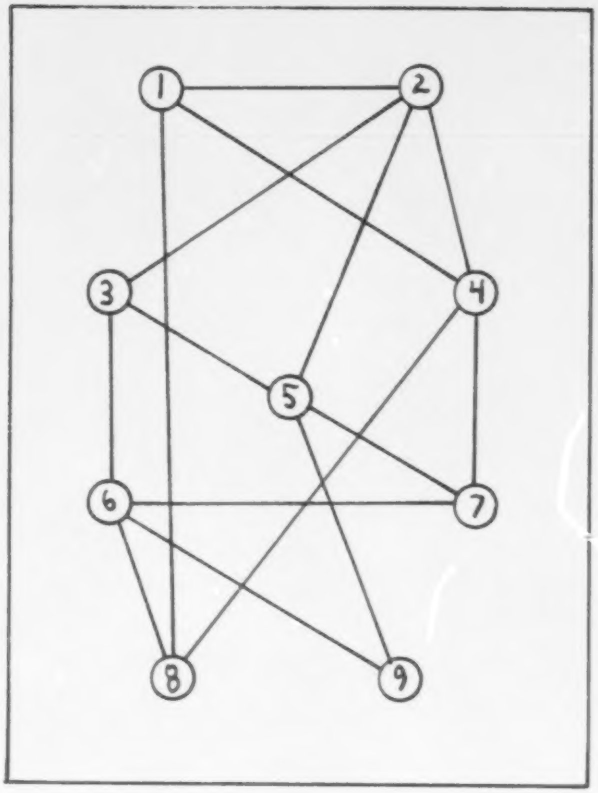
Figure 10

THE DISTANCE MATRIX AND CRITICAL NODES

On the upper left-hand corner of this figure is shown a hypothetical nine-member network. To the right of this is the distance matrix for that network. The rightmost column of the matrix contains the means of the rows of the matrix. The values in this column are thus the mean number of steps it takes that node to reach all other nodes. The overall mean for the group, together with the standard deviation of the distribution of means, is shown below the matrix.

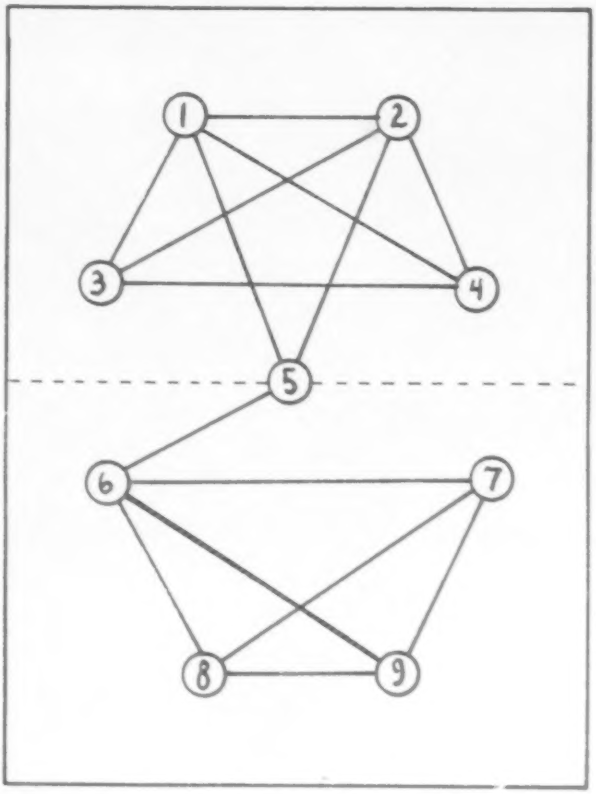
The network in the bottom left-hand corner is an example of the kind of situation that occurs when two or more groups are identified as a single group. Clearly, node 5 is a liaison between the two groups. The middle matrix on the right half of the page is the distance matrix for this group. Note the relatively high standard deviation for this group, compared to the one above it.

The third matrix was constructed after removing node 5. Note that there are no values for many of the elements, indicating that the group is no longer connected. The means shown for this bottom matrix are the values that would be obtained if the group were split in two, and the means for each group calculated separately.



	1	2	3	4	5	6	7	8	9	
1	0	1	2	1	2	2	2	1	3	1.75
2	1	0	1	1	1	2	2	3	2	1.62
3	2	1	0	2	1	1	2	2	2	1.62
4	1	1	2	0	2	2	1	1	3	1.62
5	2	1	1	2	0	2	1	3	1	1.62
6	2	2	1	2	2	0	1	1	1	1.5
7	2	2	2	1	1	1	0	2	2	1.62
8	1	3	2	1	3	1	2	0	2	1.87
9	3	2	2	3	1	1	2	2	0	2.0

MEAN: 1.691
S.D. = .146



	1	2	3	4	5	6	7	8	9		
1	0	1	1	1	1	2	3	3	3	1.87	
2	1	0	1	1	1	2	3	3	3	1.87	
3	1	1	0	1	2	3	4	4	4	2.5	
4	1	1	1	0	2	3	4	4	4	2.5	
5	1	1	2	2	0	1	2	2	2	1.62	←
6	2	2	3	3	1	0	1	1	1	1.75	
7	3	3	4	4	4	1	0	1	1	2.61	
8	3	3	4	4	4	1	1	0	1	2.61	
9	3	3	4	4	4	1	1	1	0	2.61	

MEAN = 2.215
S.D. = .399

	1	2	3	4	5	6	7	8	9		
1	0	1	1	1	-	—	—	—	—	1.0	"A"
2	1	0	1	1	-	—	—	—	—	1.0	
3	1	1	0	1	-	—	—	—	—	1.0	
4	1	1	1	0	-	—	—	—	—	1.0	
5											
6	—	—	—	—	-	0	1	1	1	1.0	"B"
7	—	—	—	—	-	1	0	1	1	1.0	
8	—	—	—	—	-	1	1	0	1	1.0	
9	—	—	—	—	-	1	1	1	0	1.0	

FIGURE 10

identified as a liaison. That node will later be tested for the criterion of group membership and if it passes, will be returned to one of the groups.

The problem has thus been reduced to one of identifying any critical nodes which may exist in a group. If there is one, it will be the node with the lowest average distance from all other nodes. This is because all paths from nodes in either half of the group to the other half must go through the critical node. The average distance from any node to all the other nodes is given by the average of all the entries in that node's row in the distance matrix. This is illustrated in Figure 10. If there is a set of critical nodes, they will be the nodes with the smallest row means.

The fact that critical nodes have lower row means than the other members suggests that there must be some variation in the row means if there are any critical nodes. We can take advantage of this fact if we only look for critical nodes when there is some variance. It turns out that this leads to a large savings in terms of computation time. This is because of the way we test for critical nodes.

To check a node to see if it is critical, we remove it from the group and re-calculate the distance matrix. If, as a result of the removal, the group becomes disconnected, we have found a critical node. If the group is still connected, we try the next candidate -- the node which, of all the remaining nodes, has the smallest row mean. We will stop this process after taking out some percentage of the original group members (usually ten percent is enough to "catch" all the critical nodes) if the group continues to remain connected. If this happenes, we put all the removed nodes back into the group.

It is easy to see that there is considerable effort involved in searching for critical nodes. This is why the heuristic device of checking the variance of the row means is so important. In every network that has been examined so far, this heuristic has worked correctly. That is, it did not prevent any critical nodes from being found. Similarly, the approach of looking at nodes with the lowest row means always finds the critical nodes. The optimum value to use as a cutting point for the variance test seems to be about 0.3. Whenever the standard deviation of the row means exceeds this value, there is likely to be a critical node. Whenever the standard deviation is less than this value, there is not.

After all groups have passed these tests, the obtained classification of nodes to groups and other roles will be exact. At this point, various indices may be calculated and the results tabled in any convenient manner. A flow chart of the algorithm is shown in Figure 11.

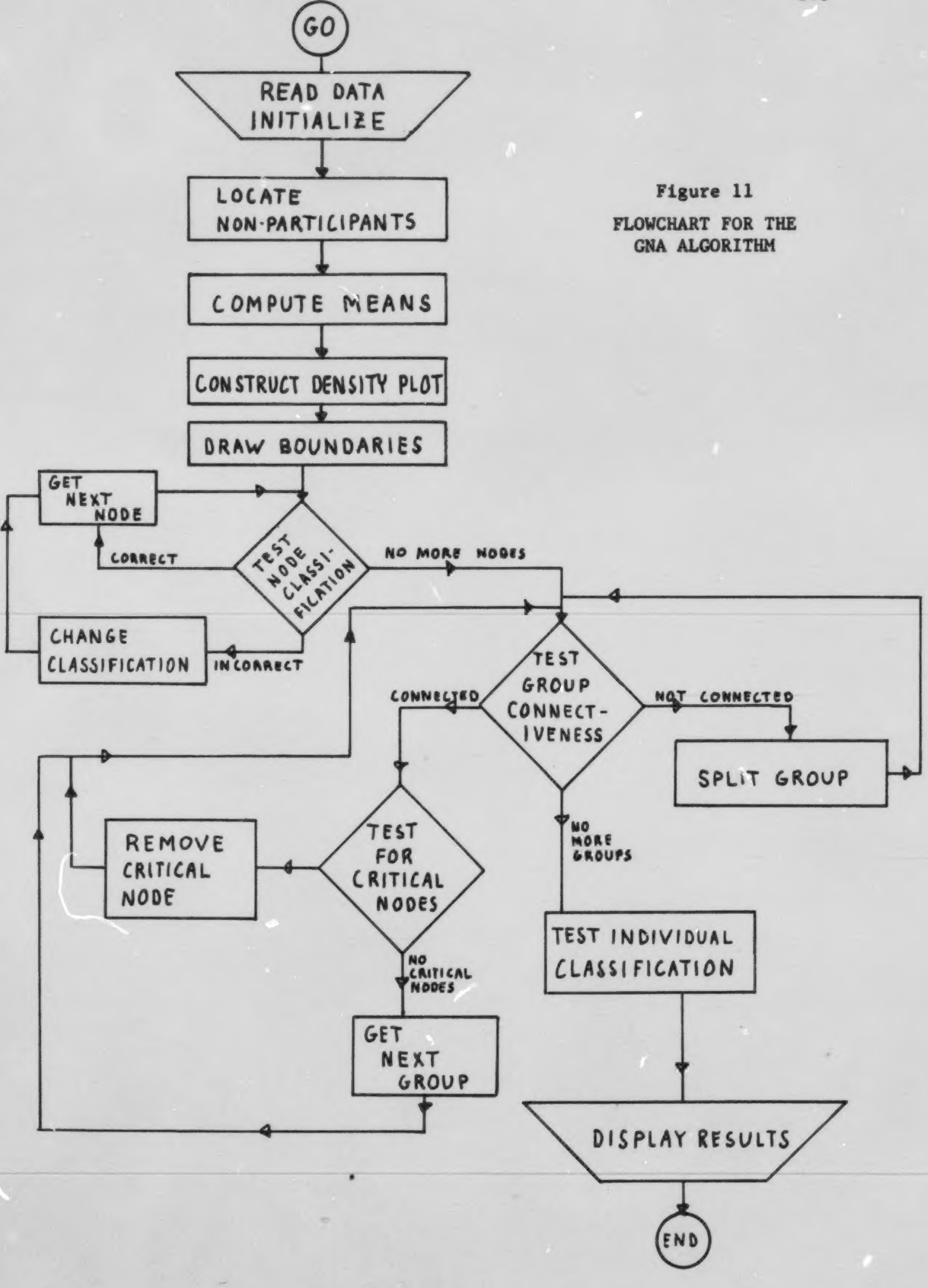

GO
READ DATA
INITIALIZE
LOCATE
NON-PARTICIPANTS
COMPUTE MEANS
CONSTRUCT DENSITY PLOT
DRAW BOUNDARIES
GET NEXT NODE
TEST NODE CLASSIFICATION
CORRECT
INCORRECT
CHANGE CLASSIFICATION
NO MORE NODES
TEST GROUP CONNECTIVENESS
CONNECTED
NOT CONNECTED
SPLIT GROUP
NO MORE GROUPS
TEST FOR CRITICAL NODES
REMOVE CRITICAL NODE
NO CRITICAL NODES
GET NEXT GROUP
TEST INDIVIDUAL CLASSIFICATION
DISPLAY RESULTS
END

Figure 11

FLOWCHART FOR THE
GNA ALGORITHM

CHAPTER NINE

THE SYSTEMS APPROACH, GASSICS, AND GNA

The analytic method introduced here can be viewed from three levels of specificity. First we have a general systems approach, which specifies both some general characteristics of the objects we are interested in and the processes by which we become informed about those characteristics. Second we have GASSICS — an analytic procedure which specifies what kinds of data are needed and what kinds of operations ought to be done on the data; and finally we have the three stages of the GASSICS procedure, in which explicit and complete operationalizations are specified.

The intent of this chapter is to examine the analytic method presented in Part Three at the three levels of specificity and see both what is included at each level and how the levels relate to one another. Then we will relate several different aspects of the approach described here to other analytic methods. A large part of the discussion here will be concerned with the comparisons of specific techniques of analysis in the context of the more general GASSICS procedure.

THREE LEVELS OF SPECIFICITY

In this section we show what is included at each level and how the levels are related to one another.

The General Systems Approach

At the lowest level of specificity we have a concern for basic, fundamental issues: What are we studying? How do the processes involved in "gaining an understanding" work? We started the explication in Chapter Three with an analysis of the form of systems. There we outlined the conceptual model we use to organize our ideas about structure.

The central concepts included:

- interactions, viewed as constraint;
- levels, as concepts of description and thus distinction;
- emergent versus additive properties; and
- logical restrictions on interactions between systems.

When we combined these ideas with the processes of observation and description, we obtained a general framework for approaching complex systems.

This general framework suggests that we respect the multiple-leveled complexity of systems by approaching them in ways that do not oversimplify them. That is, research methods should be designed to work in situations involving multiple, rather than single, levels.

Furthermore, the framework suggests that the most fruitful places to look for an understanding of emergent behaviors will be in the relations between levels of the system. Again, this is a multiple-leveled approach.

Much previous communication research has been criticized on the grounds that follow.

(1) It has a "psychological bias." Since the unit of response is the individual person, the person has been the unit of analysis in most communication research. This approach completely ignores all system effects, which can only be seen in the larger system. Not only are system effects ignored, but also the very existence of the system is not considered. It is as if the concept of system were completely foreign to people working in the field (see Rogers, 1975).

(2) While communication is characterized as a process, most communication research is not process oriented. Dependent variables are isolated and measured at one point in time. Most of the variables that have been studied are effects variables. This approach, then,

effectively treats communication as a "black box" which is studied only indirectly, by analysis of its effects.

(3) The conceptualizations and operationalization used in most communication research demands one-way causal relationships. There are independent variables (causes) and dependent variables (effects). This is, of course, the classical analytic approach, which denies the existence of mutual causal loops. Such loops do not fit into the model either conceptually or operationally.

The general systems approach, as it is presented here, answers these objections.

(1) There is no psychological bias, because systems are viewed both conceptually and operationally as having multiple levels. Although analysis may be at the same level from which the data are collected, it can also be at several different levels. Thus, we collect data describing behaviors of individual persons and make statements about the whole system and groups of people working together, as well as about the individual people. When speaking of the groups and other higher level units, we speak of them as unified wholes, rather than as collections of individuals. That is, we speak of the properties of the higher level units as units, rather than the properties of the individuals making up the units.

(2) In a multiple-leveled systems approach, any given phenomenon will look different when viewed from different levels. At one level we may have a dynamic behavior that is observed only as changes over time. At a higher level, this same behavior may appear to be a structural characteristic of the system. Thus, process at one level becomes structure at another. The

systems approach respects the processual nature of communication phenomena by starting with data descriptive of relationships -- dynamic processes which occur across time. Again, it is the multiple-leveled paradigm that suggests how this is done.

(3) The systems approach does not divide phenomena into discrete causes and effects. Indeed, some "variables" may function as both "causes" and "effects" when they are parts of mutual causal loops. There is a whole set of systems concepts that have been developed explicitly for these situations. These concepts, including both positive and negative feedback, are the basis for the science of control -- cybernetics. Although the ideas of cybernetics are not discussed in this dissertation, they are compatible with the approach outlined here.

The Second Level of Specificity: The GASSICS Procedure

In Chapters Four and Five we discussed the implications of there being multiple levels on the processes of observation and description. One conclusion we made there was that the first step in the investigation of an unknown system would be to elucidate the organization of the system -- how many levels are there and what is the structure at each level. The GASSICS procedure, outlined at the end of Chapter Five and discussed in detail in the chapters of Part Three, performs this structural analysis. Once the organization of the system is clarified in this way, the dynamics of the system can be explored. Only after this entire process has been completed would we say we understood the system.

The GASSICS procedure, then, is only one step in the analysis of systems, and GNA is only one step in the GASSICS procedure.

The GASSICS procedure is an iterative loop that makes available information about the organization of the system at multiple levels. Because of the differences between levels and the relationships between phenomena at different levels, the procedure progresses one level at a time. The complete application of the procedure, with one cycle for each higher level, will be a multiple-level analysis. However, during any single cycle of the loop, a maximum of two levels will be under consideration.

The key idea of the GASSICS procedure is the repetitive looping. Within each cycle a sequence of operations is performed. The sequence of operations in a cycle is ordered so that the end results of the cycle are fed back into the beginning parts of the next cycle. This "feeding back" process is where the new (alternate) descriptions are formed. It is thus the strategic move that makes it possible to examine multiple-leveled systems with analytic tools that are not unreasonably complex.

Where the adoption of a systems approach specifies a set of conceptual "biases" that later determine operational procedures, the adoption of the GASSICS approach specifies the sequence of operations that must be performed in the analysis. It does not, however, specify how those operations are to be performed.

The Highest Level of Specificity: The Actual Operations

At the highest level of specificity we have the actual operations that are performed in a loop of the GASSICS phase of system analysis. Several things happen when we move to this higher level of specificity. First, as specificity goes up, generality goes down. In any of the stages of the GASSICS loop we are concerned only with a limited subset of aspects related to the system. The organization of the entire analytic approach is thus hierarchical. A parallel sort of hierarchical problem-solving is seen in day-to-day existence. For example, if going shopping is parallel to system analysis, driving to the store and purchasing food might be parallel to the GASSICS procedure and analysis of the dynamics of the system. Getting in the car, starting the engine, and so on, would parallel the operational procedures that make up the steps of GASSICS.

If we have decided to use the GASSICS approach, we are still partly free to decide how to carry out the operations specified at each step. The remaining freedom is evident, for example, in the third step -- the identification of any differentiated units. In Chapter Two we discussed a number of methods which might seem suitable for this purpose. Given that there were already a number of methods available to do this type of analysis, it might seem perverse to develop yet another method, as we did in Chapter Eight. Why was this done?

The other methods were judged to be not suitable for one or both of two reasons: either (a) the method did not fit conceptually, or (b) it did not work well.

Rather than using a mismatched method or trying to fix one that didn't work well, we started over. This was advantageous for different reasons. Since we knew the restrictions we were faced with as well as the desired analytic goals, we could start with a clear conceptual foundation. Right away a number of alternative operational approaches were eliminated. For example, we chose a linkage method rather than a distance method. Also, knowing what we were looking for in terms of groups, we could design a method that would find just exactly what we wanted, rather than decide that we should want what the method produces. Thus, with GNA we look for groups with a particular set of characteristics. When other methods, such as factor analysis, do not include these characteristics, we decide that the methods are not matched to our goals. The development of our analytic techniques followed the formulation of conceptual goals.

The problem of getting a method that worked well was not so easy to solve as the problem of getting one that did what we wanted it to do. One thing that helped enormously was having pre-defined goals -- this gave us a standard to use in judging the efficiency of our methods. We knew what we wanted. If the use of a method did not produce the desired kinds of results, that method could be rejected. If, on the other hand, the method did work, we knew we were on the right track. To show how this developmental process worked, we discuss chronologically the history of the methods used in each of the steps of the GASSICS procedure in the next section.

THE DEVELOPMENT OF THE ANALYSIS PROCEDURES

Although the network analysis step of GASSICS comes last in the loop, it was the first of the three steps to be developed into a sophisticated tool. At the time it was being developed, data gathering methods were more advanced than data analysis methods, even though they were both primitive, compared to what we use now. The problem of structural analysis had not even been considered, and it was three or four years until the idea became useful.

G-Network Analysis: Four Breakthroughs

In late 1970 I became aware of the fact that there was no fast, easy way to perform network analysis. The method that was used most for large systems was matrix manipulation. In 1970, systems having over 200 members were considered to be large. I started to work on a computerized algorithm that would perform the matrix reordering task.

There were four critical breakthroughs during the development of the current version of the GNA algorithm. These breakthroughs made it possible to program a computer to do the work of identifying the groups in a very large network. They are discussed in the next four sub-sections.

1. The First Breakthrough: NETWOW

The earliest version of the current technique is described in Richards (1971). The first working computer implementation of the method, NETWOW, did essentially what the matrix manipulators (e.g., Weiss, Jacobson, and Seashore, etc.) were doing by hand — it reordered the rows and columns of the adjacency matrix so that non-zero entires clustered about the main diagonal. Although the results were similar to those obtained by the earlier investigators, the process by which these results were obtained was different from the ones they used.

NETWOW, like NEGOPY, used an iterative procedure which treated links as vectors. In each iteration of NETWOW, means were calculated just as they are calculated in NEGOPY. In NETWOW, however, after the means were calculated they were rank-ordered. The node with the smallest mean was put in row 1; the node with the next-smallest mean went in row 2, and so on. This process was repeated ten or twelve times. Although it seemed strange to take means of subject numbers, the process worked. Nodes that were linked to each other tended to be placed close to each other in the matrix. The discovery of this simple procedure was the first breakthrough in the development of an efficient method for GNA.

The output of NETWOW was the reordered matrix. Groups were identified by hand inspection.

A major problem with NETWOW was that it did not identify the groups -- it just made them easier to spot. In the summer of 1972 I began to develop an algorithm that would actually identify the groups. I started with the method suggested in Richards (1971). The method suggested there was to scan the final list of means, looking for points at which there were large breaks or gaps when going from one mean to the next. Group boundaries would be located at these breaks. This method did not work. The problem seemed to be due to the matrix representation of the data.

2. The Second Breakthrough: "Continuous Matrices"

The matrix technique, obviously, uses a matrix. Each node has a row and a column. Rows and columns are discrete: all rows and columns have the same "width."* Since rows are ordered the same way as columns, nodes can be represented as points (more correctly, segments) along the diagonal. In the process of manipulation, which makes groups visible, rows and columns are simultaneously permuted. This means that their order is changed. Only the order changes — the widths of rows and columns remain the same.

This is where the sociometric use of matrices diverges from the mathematical use. In matrix manipulation, groups are visible as blocks of non-zero elements along the diagonal. This is a graphical interpretation of a numerical representation of the data. This graphical interpretation would work a lot better if the representation were altered slightly so that it would be a graphical representation, instead of a numerical one. The alteration required is a very simple one. The "location" of nodes in the "matrix" has to be changed from the discrete ordinal representation of mathematical numerical matrix style to an interval or ratio level representation of graphical style. Let us make this clearer.

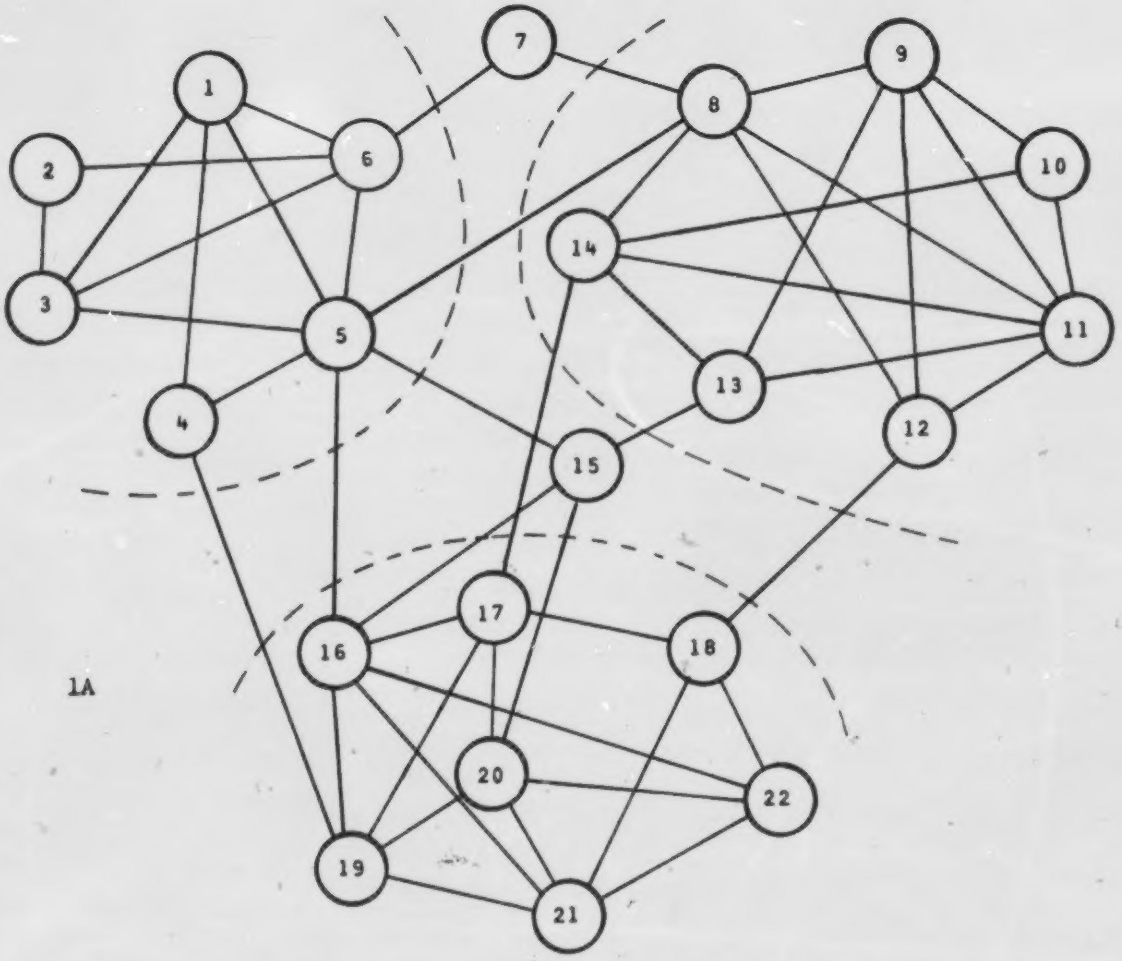
---

*The "width" is not really defined, until the matrix is graphically displayed. In practice, the rows and columns of ink-on-paper versions of matrices are of equal widths. This is a matter of convenience, as the mathematical concept of a matrix is concerned only with the actual numbers in the rows and columns, and not with the way they are written.

We will call the numerical mathematical matrix with rows and columns of unit width a "discrete" matrix. In a "discrete" matrix, nodes are "located" at row 1 or row 2 or row 3, etc.; and never at row 1.3 or row 2.9 and so on. Nodes are all one unit away from their neighbors. "Location" is ordinal only, and order is discrete.

In the other kind of matrix this is not the case. Call this other situation a "continuous" matrix. It is not really a regular matrix, but the analogy works. In a "continuous" matrix, it is as if the rows and columns can have varying "widths." Thus there can be nodes at 1.0, 1.2, 1.3, 6.7, and any other place along a continuum. If two nodes are very close together, they can be put very close together. In the "discrete" matrix, they must always be at least one unit apart. Location in a "continuous" matrix is an interval value. Although locations can be reduced to orders, doing so throws away information. A "continuous" matrix is shown in comparison with a "discrete" one in Figure 1.

There are at least two important differences that result from the shift from "discrete" to "continuous." The first is the efficiency of the rearranging process, which used to be called the "reordering" process. The NETWOW program used a "discrete" matrix. Means were calculated in much the same way that they are calculated in the version of GNA described in Chapter 8. In order to locate the nodes in the matrix after the calculation, the means were rank-ordered, and nodes assigned to rows and columns on the basis of their ranking. The next round of calculations was done using the rank orders -- the new row and column numbers (see Richards, 1971).



1A

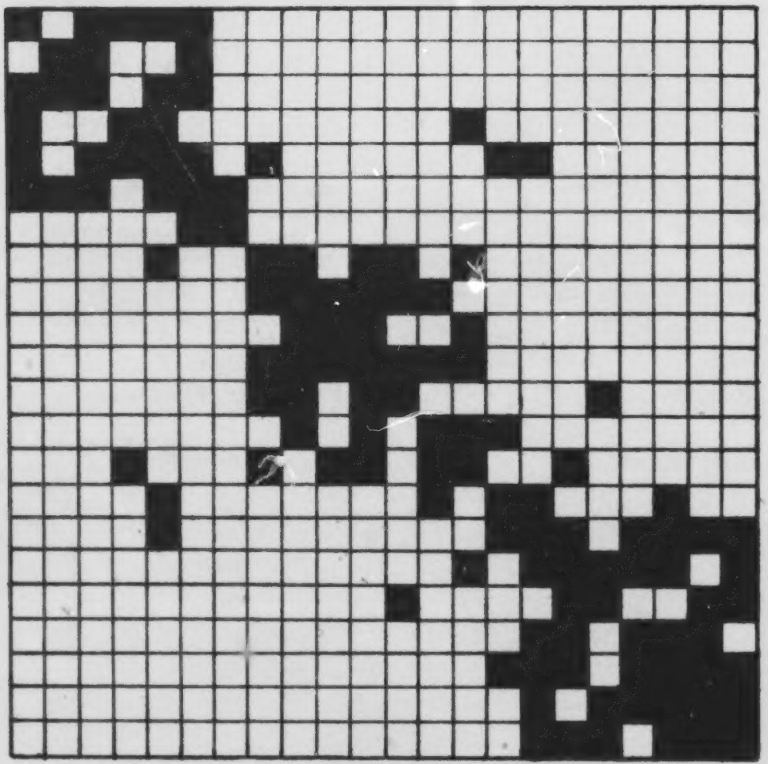
Figure 1

In 1A above, a hypothetical network with three groups is shown. 1B on the next page shows a "discrete" binary adjacency matrix for the network in 1A. Note that the groups are visible as clusters about the diagonal.

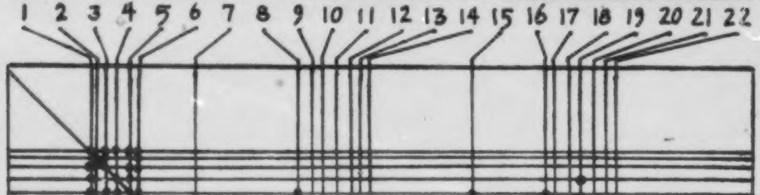
1C, on the bottom of the next page, shows a "continous matrix" that might be the result of rearranging the rows and columns of the "discrete" matrix in B. Note that in going from B to C, only the "widths" of columns and rows has changed. In this "continous" version, groups are clearly visible as dense clusters along the "diagonal" of the "matrix."

The "diagonal" of the "continuous" matrix is projected onto the line segment shown in 1D, below C. This line segment becomes the "continuum" spoken of in the text.

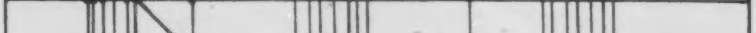
1B



1C


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

1D



In the "continuous" representation, nodes are not assigned to discrete rows and columns, but rather to infinitely small points that can be anywhere on a continuum. This is the crucial difference. With the "discrete" matrix, a lot of information is thrown away with each rank-ordering, as actual numbers are replaced by ordinal ranks. With the "continuous" form, there is no rank-ordering. Nodes are "moved" to the points indicated by their means, rather than the points indicated by the ranks of their means. As a result, the additional information is kept and used. This makes the whole rearranging process work much faster -- converging to a relatively stable configuration in four iterations instead of ten or twelve. This difference in efficiency is directly due to the shift from "discrete" to "continuous."

The second difference resulting from the shift is in the "readability" of the final representation, which can also be seen in Figure 1. In B is shown a "discrete" matrix for a sample network. The "matrix" shown in C is a "continuous" one. Because rows and columns all have to be exactly one unit wide in the "discrete" version of B, it is much harder to see the groups there than in the "continuous" version of C. This is the second difference due to the shift from "discrete" to "continuous." This shift was the second breakthrough; it made it possible to begin work on the algorithm that would actually identify the groups.

It is very important to keep one thing about these "continuous matrices" clear: the "location" of nodes in these "matrices" is really quite meaningless as a description of the system. Because of the iterative procedure which is used to obtain the final "locations," it is not possible to say exactly what the relation between original data and final result is. The final "location" is, however, a valuable heuristic that, when properly interpreted, makes the task of delineating groups much easier than it would otherwise be. In the version of the GNA program described in Chapter 8, the "matrices" are never drawn out. In fact, they don't really exist at all. They are only used as a way of explaining how the program works with the data.

The most that can be said about them is that they are used as clues to give a good guess as to where the boundaries of the groups lie.

3. The Third Breakthrough: The Group Identification Algorithm

While the first two major developments, vector averaging and "continuous matrix" representation, were fairly "sharp," the third, the initial group detection routine, was much more gradual. I began with the idea of locating groups on the basis of the way nodes clustered together in the "continuous matrix." The process of identifying groups from the matrix is a process of "digesting" the information in the matrix. For a starting point, it seemed to be a good idea to look for areas in the matrix that were "densely packed" with nodes. Operationally, it wasn't necessary to have the whole matrix present, since rows and columns were in the same order. This meant that a single line segment, along which nodes would be located, would be sufficient. Thus, I started to work with the "continuum" described in Chapter Eight. As is shown in the bottom of Figure 1, the continuum is closely related to the imaginary "continuous matrix" shown in C. It is, in fact, operationally equivalent to the diagonal of that "matrix."

The task was then to locate the areas of high density along the continuum. Because the computer can only consider one thing at a time, it cannot be told to "look at all the points on the continuum, and pkck out the areas of high density." Instead, the density at each point has to be calculated, and then the densities are examined.

The first step was thus to devise a measure of density. I tried a succession of methods here. They all used a "window," which I would slide down the continuum. I would note the number of nodes visible through the window at each point and use this number as an indicator of density.

The first question I had to deal with was how to move the window. At how many points should measurements be made? Which points would be the best to measure from? There was no logical basis to suggest which method should be chosen.

I took a practical approach. I constructed a data set for a network with a known group structure. After computing the values for this network's continuum, I tried several methods of moving the window. This allowed me to see how each method worked, and to pick the best one. The methods that did not work included centering the window at points one unit apart and centering it at points one tenth of a unit apart. In fact, all methods that moved the window smoothly along the continuum were inferior to one that centered it on successive nodes. This is the method that was adopted and the one described in Chapter Eight.

The next step was to analyze the density plot -- the result of the scanning of the continuum. The initial idea had been that groups would appear as clearly identifiable "mounds" on the density plot. Gaps between mounds would signal the boundaries of groups. Although this simple method worked in a few very clear cut cases, where there were no links between members of groups and anyone not in their groups, it was not dependable. The problem was that the window had to be so wide that the density would not drop very far between groups. Instead of there being clean breaks, there were only slight depressions. This is shown in Figure 2. The cause of the problem was that at the end of a group, several nodes would be visible through the window. These would be the nodes just passed earlier in the group. At the beginning of the next group several nodes would also be visible -- the nodes later in the group. Thus, the density would not drop through the transition from one group to the next.

If the width of the window was descreased, the density would go to zero between groups, but it would also go to zero in the middle of some groups. The solution to the problem was based on the fact that although the number of nodes might not change when going from one group to the next, the nodes themselves would be different. Instead of being nodes already past, they would be nodes yet to come. That is, when leaving group A, the nodes being counted were members of group A, and were in the left half of the window. When entering group B, the nodes being counted were members of group B, and were in

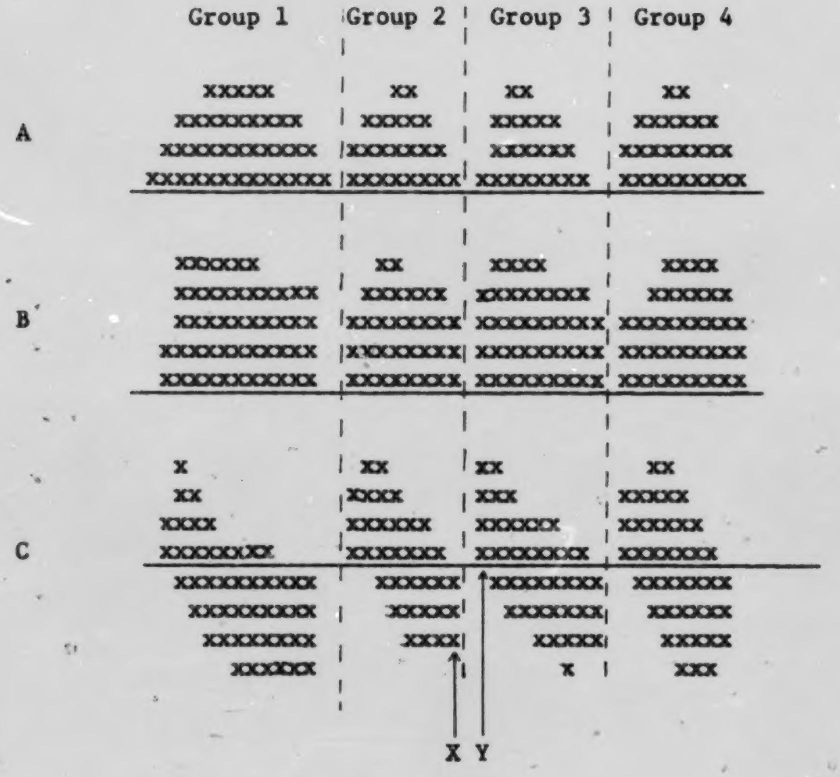


Figure 2

DIFFERENT WAYS OF CONSTRUCTING THE DENSITY PLOT

In A is shown what was expected of the density plot. Gruops are clearly separated by distinct breaks in the plot.

B shows the density plot actually obtained by plotting the number of nodes visible through the window. (Compare this plot to the one shown in Figure 7 of Chapter 8.) Note the gradual depressions between the mounds of the plot. It is difficult to say exactly where one group ends and the next begins.

In C a plot obtained by counting separately the number of nodes visible in each half of the window is shown. Note the extreme clarity of the groups, as compared to B above. Note also the large shifts between groups, as in the transition from Group 2 to Group 3 at points X and Y.

the right half of the window. In the middle of the group, visible nodes would all be from that group, and would be on both sides of the window. This was the basis for the shift to the method of separately counting the number of nodes visible in each half of the window. The result of this shift is shown in Figure 2 in this chapter, and also in Figure 7 in Chapter Eight.

The final step in the group detection phase is to draw the actual boundaries of groups. I had originally thought that gaps between groups would be clearly visible and useful as indicators of group boundaries. The problems with this assumption were already discussed. It was possible, however, to take advantage of the extra information contained in plots like the one shown in Figure 2C. Here, it is not only the low densities that mark group boundaries, but also a shift in the set of nodes being counted. This appears on the plot as a shift from several X's below the line to several above the line. This can be seen at points x and y on the plot in Figure 2C. The method of drawing group boundaries described in Chapter Eight was based on this approach. I regard the development of the group detection routine as the third breakthrough in the work on the current GNA algorithm.

4. The Fourth Breakthrough: The Exact Criteria

At this point I began to test the algorithm with real data. I discovered that the "groups" the method identified were really quite poor guesses. For example, in many cases there would be nodes imbedded in the middle of groups when there were not connections between the node and the members of the group. In other cases, there would be two or three groups all lumped together into one cluster. Sometimes, one group would appear as two or even three clusters.

These problems were all due to the inelegance of the method that was used to obtain the groupings. The source of the problem is in the statistical behavior of the mean used in the vector averaging process. Two entirely different sets of numbers can have the same mean.

In order to compensate for these difficulties, it was necessary to develop some strict logical tests that would locate incorrect groupings and correct them as they were found. This was the spur that led to what I consider to be the fourth breakthrough -- the development of both the criteria for group membership and the exact specification of network roles.*

The analytic procedure thus has two phases -- an inexact, "tentative" one and an exact final one. The first phase includes both the interative vector-averaging process and the window-scanning density-analysis group-detection process. The second phase involves the application of the strict logical tests of the various criteria for group membership and other network roles.

The two phases cannot be used separately, since each performs only part of the analysis. The first produces results which can be used as inputs to the second, but which are not useful for any other purpose. The second will not work without the information produced by the first. Together, the two phases make up a powerful procedure which extracts much of the information contained in the structure of the system being analyzed.

Data Collection Methods: All Links Are Not Equal

The development of the techniques used in the other two steps of the GASSICS loop took an entirely different path. Data collection methods started rather crudely, with instrumental questions like "Who do you talk with?" This simple question was soon modified by the addition of content qualifiers -- "Who do you talk with about job-related matters?" When it was realized that greater precision in measurement would (hopefully) lead to better results, the question was changed from the simple binary we-do-or-we-do-not-talk to a question that allowed several levels of interaction to be coded. Thus, respondents could indicate that they talked once a month or less, once a week or less, and so on.

1. Nominal Binary Approaches

Throughout the 60s and early 70s, the additional frequency information was used mainly to give the investigator some extra options. Even though more than simple binary information was available, most analyses were done with data that had been reduced to binary form. Investigators would typically do an analysis using only links that occurred once a week or more, or once a day or more, and so on. With this scheme, then, links above the chosen cutting point were kept and links below were rejected. All links that were kept were treated as if they had the same strength.

---

*See Chapter Eight.

2. Greater Precision: Ordinal Strength Indicators

In 1973 I modified the GNA algorithm and the NEGOPY program so that they would take link strength into account. Thus, the present version works on the basis of amount of linkage, rather than simple number of links. With this version, for example, it would take two links, each with a strength of 2, to be the equivalent of one with a strength of 4. This approach, therefore, retains and uses the extra information that is thrown away when the data are reduced to simple binary form.

Extra precision was added in yet another extension of the "all links are not equal" philosophy when the algorithm and computer program were modified to consider not one but two indicators of link strength. The first measurement instruments to incorporate this feature operationzlied the second indicator as importance. In the analysis stage frequency and importance would be combined in some predetermined way by the program to give a single index of link strength.

3. Still More Precision: Ratio-level Indicators

The final modification in the direction of greater precision was first used by Monge and his colleagues at San Jose State University. They replaced the discrete category system for coding frequency of interaction with a continuous indicator. There are two versions of continuous frequency reporting. The first asks "How many times?" in some standard time period the respondent interacts with others; the second asks for estimates of total duration of interaction -- "How many hours do you spend in a typical month interacting with __________?"

Since both continuous measures are direct ratio estimates of interaction time, they can be used in analysis as they are, with no additional transformations.

At the time of this writing, none of the methods have been tested empirically for reliability or cross-method consistency.

Structural Indicators

The idea of structural measures was first conceived after a large number of networks had been analyzed using the NEGOPY program. With many datasets, the program would not be able to break the system down into groups, unless either weak links or unreciprocated links were excluded from the analysis. (This suggests that weak and unreciprocated links are less reliable than stronger, reciprocated ones.) With many datasets, then, it was necessary to submit several runs before the results would be useful.

It seemed that, if a heuristic measure of structuring would be applied to the data before any network runs were done, it would be possible to predict whether or not there would be any differentiation into groups. This would save a lot of time and money, as network analysis would only be done on data that warranted the effort. This was the spur that led to the development of the structural measures described in Chapter Seven.

It would be good to be able to report that the structural measures have been used and found to work as expected in a variety of situations. Such a report is possible, but with qualifications. The main limitation is that the statistics were worked out only recently (see Richards, 1974), and programmed even more recently (within the last two months). Four networks were analyzed using this version of the program and the measures did work as expected.

One of these four datasets was run three times, with successively stricter limits on weak and unreciprocated links. As more and more of the less reliable links were excluded from analysis, the structural measure reached higher and higher values, as would be expected. It is interesting and significant to note that separate measures of structure -- the percentage of links explained by group boundaries and group densities -- both increased in parallel with the structural measure in this manipulation. Therefore, the measure proposed in Chapter Seven appears to work as expected, at the same time that it is supported by other independent measures of the same general concept.

Although the GNA technique, as implemented in the NEGOPY program, is embedded in the GASSICS procedure, it has not yet been used in that context. The most complete studies to date have only gone one loop through the cycle, with the structural calculations coming last, rather than second. It seems that, as the structural indicators become better understood, they will become a more integrated part of network/systems studies.

Finally, as the remaining theoretical and methodological problems (discussed in Chapter Eleven) are solved, there will be studies done using two or three cycles of the GASSICS loop.

In the next chapter we present an example of a study using one complete cycle of the GASSICS loop.

PART IV: RESEARCH

CHAPTER TEN

USING G-NETWORK ANALYSIS IN RESEARCH

The goals of this chapter are to give the reader some idea of the kinds of systems that have been studied, to indicate in detail the structural characteristics of a fairly typical network, and to show how some researchers have attempted to relate network variables to other phenomena. We do not attempt to test any hypotheses. It seems that a familiarity with the kinds of things that are found over and over again in networks is necessary before theory is advanced and tested.

Putting the description of the example network into the context provided by the analysis of several other communication networks enhances its value as a representative example. This particular network was chosen both because it was representative and because it was "well-behaved." Both of these characteristics are desirable in an example.

THE ORGANIZATION OF THE CHAPTER

The rest of the chapter is divided into three sections. In the first, several communication network analysis projects are reviewed in order to give some idea of the uses to which the NEGOPY program has been put. A second function of the review is to give background information about the networks, some of which are used in cross-network comparisons in later sections.

In the second part, a medium-sized communication network (the Melbourne network) is examined using the techniques described in the chapters of Part Three of this dissertation. An examination of the amount of structuring of the system, using the concepts and operations described in Chapter Seven, begins the analysis. Next comes an examination of the organization of the system, in the terms outlined in Chapter Eight. Here we identify the subsystem structure of the network by applying the GNA algorithm to the data. The results are then examined from a systems poknt of view, which stresses the importance of the relationships among the parts at each of several levels of analysis.

An important part of the analysis in this section is the comparison of the results from the Melbourne study to the results obtained in several of the networks described in the first section. With these comparisons it is possible to get some idea of the kinds of patterns that are repeated in a variety of systems of different types and different sizes.

The third section discusses some examples of work that has been done in relating network variables to other phenomena. The studies examined here are a subset of the ones used in the earlier comparisons.

REVIEW OF SELECTED NETWORKS

Several networks have been studied with the NEGOPY program since it was first developed in the early 1970s. Brief descriptions of some of these studies make up the rest of this section. A "*" by the description head indicates that data from the study are used in the cross-network comparisons in the next section.

1. The Chase Studies*

Three datasets were gathered at different time periods in different units of Chase Manhattan Bank. Although these studies were done mainly to provide information to the management of the Bank, they are of interest to the scientific community at large. The results of these studies have not been published; however, MacDonald (1970) has written a manuscript which discusses the application of network theory in the organizational context. Some of the characteristics of the largest of the three networks are also summarized by Farace and Johnson (1974). In this study, persons were asked to indicate how often they talked with other people in their part of the

Bank. Communication was broken down into three content areas: Production, Innovation, and Maintenance.

Production was defined as communication related to getting the job done. This category would be conceptually close to the traditional ideas of "formal" communication. Innovation communication was concerned with changes in organizational procedure and new ways of doing things. Maintenance communication is devoted to social maintenance functions and is conceptually similar to the "informal" communication often seen in the literature.

Because of the size of the system (N=973), it was felt that a roster format (see Chapter Six) would be too long. Data were therefore gathered with open-ended recall techniques. Because of difficulties unique to the recall method (i.e., forgotten names, failure to remember all the communication contacts, etc.), there were many difficulties with the data. The number of isolates was high in this study (probably at least partly due to the use of the recall method) and the level of reciprocation was very low. Nonetheless, the study provides some valuable insights to the nature of communication networks in very large systems.

2. The OCD Studies*

Two datasets were collected a year apart in the Office for Civil Defense in the Pentagon. This study used the same method for data collection that was used in the Chase Studies. Here again, the Production/Innovation/Maintenance breakdown was used. The results are reported in an unpublished manuscript by Berlo, Farace, Monge, Betty, and Danowski (1972), and summarized briefly in Farace and Johnson (1974).

3. The Treasure Island Studies*

Working as an Office of Naval Research contractor, Monge has been studying the development of communication structure in military organizations. His data come from a naval base located on Treasure Island in the San Francisco Bay. Together with his colleagues, Monge has been attempting to develop and validate causal models relating several organizational variables to one another. One subset of these variables contains indices of several network characteristics at different levels of analysis. Another important aspect of this work is the use of more sophisticated approaches to the measurement of link strength: Monge was the first to use absolute number of interactions and then total interaction duration, in place of a simple frequency category system. This work is reported in Monge, Kirste, and Edwards (1974). One of the networks from this study is used in the comparisons in the next section.

4. The Naval Studies

Two goals were identified for the research program of Roberts and O'Reilly: (1) " ... to find and adopt a general framework to guide research in organizational communication ... " and (2) " ... to begin to identify communication variables within each observational level which might be inter-related ultimately and then related to other important organizational behaviors within each level. Across level relationships might then be explored profitably" (1975 a, p. 2). Summing up their theoretical rationale, Roberts and O'Reilly say:

> Looking back to our framework, and the criteria adopted for specifying social systems it was now necessary to define communication structures which might later be related to processes; such as perceptions about communication dysfunctions; the relationships of those perceptions to other responses such as job satisfaction and performance; and the relationship of communication processes to characteristics people bring with them to their jobs, such as personality, etc. It was also necessary to examine shifts and stabilities in structure over time in order to identify organizational regularities and dynamics ... it was felt that communication structures should be identified at the individual, group, and organizational levels, before returning to each level to investigate relationships of structures to antecedents and to process responses. (1975a, p. 6)

The respondents in this study were officers and enlisted personnel in three high technology military units. Data were collected three months after the units were commissioned and again one year later. In both phases respondents were given a sociometric survey which was designed to tap several networks: (1) an expertise network -- "when you need technical advice in doing your job who are the persons you are most likely to ask?"; (2) a social network -- "with which persons in this squadron are you most likely to have social conversations (not work related) in the course of a work day?"; (3) formal authority -- "if you are upset about something related to the Navy or to your job, to whom in the squadron are you most likely to express your dissatisfaction (gripe) formally?" (Roberts and O'Reilly, 1975a, p. 9).

Respondents provided the name or a description (later turned into a name) of the relevant persons appropriate to each question. Both communication frequency and communication importance were indicated for each contact. Roberts and O'Reilly (1975a, p. 10)mention that the recall technique they used cause some problems. They suggest that a roster-type instrument would give better results, but that this approach could not be used because of the size of their organization.

Data from the sociometric questions were submitted to the NEGOPY program for analysis. Results were examined for three kinds of information: (1) descriptive communication structure — what do the communication networks look like; (2) relationships among the various structural characteristics; and (3) relationships between network/structural variables and other behavioral and attitudinal variables. Some of their findings are described in the last section of this chapter.

Discussion

Fast, reliable methods of performing network analysis are new. There has not yet been enough time to develop complex models relating network variables to other phenomena. Before this more advanced work is even begun, it is necessary to become familiar with communication networks themselves, independent of other considerations. We agree with Roberts and O'Reilly when they say " ... communication structures should be identified at the individual, group, and organizational levels, before returning to each level to investigate relationships of structures to antecedents and to process responses" (1975a, p.6, italics added).

In the next section of this chapter, we focus on communication structures at the different levels suggested both by our theoretical perspective (Chapters Three, Four, and Five) and by Roberts and O'Reilly, among others.

One network analysis is clearly not enough to give us the kind of familiarity with communication networks that we would like to have. In the limited scope of this chapter, however, one network analysis is almost too much, as the details seem to defy attempts to gain a gestalt understanding of what the system is like. The only response we can offer is that it must be realized that we are working with complex multi-leveled information processing systems, and not a single independent variable-dependent variable pair. Complex systems have a disturbing tendency to be complex. It must be realized that, while parsimony and elegance have a certain attractiveness, they are not always possible.

With these comments in mind, we move on now to an example of network analysis using the NEGOPY program. We have made one concession to the call for meaning -- we have attempted to put the results of the sample network (the Melbourne study) into the context provided by several other analyses. After the discussion of structural characteristics in the next section, we attempt to give some idea of the kind of relationships between network variables and other variables that have been examined. The descriptions in this last section are just a sampling of the several studies that have been done; however, they should serve to give some idea of the nature of this research.

AN EXAMPLE OF NETWORK ANALYSIS: THE MELBOURNE STUDY

1. Background

In June 1973, the Director of the Deparmtnet of Agriculture in Victoria, Australia, authorized a study of the communication networks that were operating within the department. This was done in order to "determine the actual patterns of communication that occurred between the officers who were principally involved in decision making in the department" (Russel, 1974, p. 1). The Department had offices in Melbourne, Burnley, Werribee, Glenormiston, Horsham, Bendigo, Dookie, and Warragul -- all cities in the State of Victoria. The study involved 261 professional and administrative officers, each of which was asked to complete a questionnaire dealing with his usual communication contacts.

Each officer was asked to indicate "which officers he communicated with about each topic area, and how frequently he communicated with them" (Russel, 1974, p. 2). There were four topic areas:

(1) Administration -- communication about new and current administrative instructions, policies and procedures, finance or staff matters, and industrial relations.

(2) Technical Agriculture -- communication about the respondent's own and related fields of agriculture, economics, food technology, or veterinary science, including communication about technical aspects, interpretation, farming and farming practices.

(3) Training -- communication about training needs and opportunities for the respondent or other offices, including workshops, conferences, refresher courses, study tours, and study leave.

(4) Work Planning -- communication about planning and management of research or extension projects.

Frequency of communication was divided into the following categories:

1 = Usually once a day or more
2 = Usually once or twice a week
3 = Usually once or twice a month
4 = Usually once or twice every three months

Each respondent was provided with a list of all the officers in the study with places provided for responses for each of the four topic areas. All the respondent had to do was write in the appropriate frequency in the proper places.

The data reported here are for the *technical agriculture network*, which had the least relationship to the formal hierarchy of the organization. It was felt that this would give a better idea of how networks that are free to organize themselves are structured.

2. Structural Analysis

In this section we perform the analysis described in Chapter Seven. The procedure is straightforward, although the computational methods are complex.

(a) *N, L, and C.* We begin with the basic system parameters. There are 261 nodes and 439 links. This gives a density of 0.013, using Equation #1 (all equations are from Chapter Seven).

(b) *S_e^2, S_o^2, and the distribution of ℓ_i's.* The next step is to examine the distribution of ℓ_i's to see if there is more or less variance than we would expect with a normal distribution of links. Calculation shows an observed variance, S_o^2, of 11.09. This is greater than the expected value of 3.32, given by Equation #5. The difference gives an F of 3.340, which is significant at $p<0.01$, with degrees of freedom of 260 and 260. Because the variance is greater than expected, we use the equations that control for the distribution of ℓ_i's in all further calculations.

(c) *$T_{e.}\ell$, $T_{m.}\ell$, and T_o.* Using Equation #7, which controls for the ℓ_i's, we calculate the expected number of triangles, $T_{e.}\ell$, to be 11. Equation #10 gives a maximum T for this distribution of ℓ_i's of 847. The observed T was 273. With Equation #12 we find this gives a relative structuring of 0.313, where 0.0 is random chaos and 1.0 is total constraint. (The values described here are shown in Table 1.) We conclude that the system is non-random. Therefore, the network will be examined for differentiated parts.

3. The Network Analysis

A 1974 version of the GNA program was used to analyze the data reported here. The control parameters of the program were set as follows:

Table 1

STRUCTURAL ANALYSIS OF THE MELBOURNE NETWORK

	EQN#**	Value	Description
L		439	Number of links
$E\ell_i=2L$		878	Number of directed links
N		261	Number of nodes
$\overline{\ell}$		3.364	Mean of ℓ_i's
C	1	$1.294\text{X}10^{-2}$	System density
C^2		$1.674\text{X}10^{-4}$	
$T_e/T_{max}=C^3$	4	$2.166\text{X}10^{-6}$	
T_e	2	6.3446	Expected T, given N and L
T_{max}	3	2929290	Maximum T, given N
S^2e	5	3.320*	Expected variance in ℓ_i, given N and L
S^2o	6	11.090*	Observed variance in ℓ_i
$T_{e.\ell}$	7	10.96	Expected T, given ℓ_i's
T_m	8	4060	Maximum T, given N and L
$T_{m.\ell}$	10	846.67	Maximum T, given ℓ_i's
T_o		273	Observed number of triangles
$n=\sqrt{2L}$		30	Minimum number of nodes that could have 431 links
S	11	0.0658	Structure, given N and L
$S_{.\ell}$	12	0.3135	Structure, given N and ℓ_i's

*Observed differs from expected, significant at $p<0.01$ (F-test, df=260 and 260)

**All equation numbers are from Chapter Seven.

- -- add links to force reciprocation
- -- consider links to be non-directed
- -- compute four iterations of vector averaging
- -- use a 2-unit-wide window for scanning in group detection
- -- use a sensitivity ratio of 100 in group detection
- -- raise raw link strength values to the third power in order to approximate ratio scaling
- -- use a 50% criterion for group membership

The Network: Basic Findings

The results of the network analysis are summarized in Table 2. Some of the numbers there should be pointed out. 77.8% of all 261 nodes were participants. This is considerably higher than the average percentage of the other networks which are compared in Table 3. 86.8% of the 190 participants were members of 28 groups.

Gross Comparisons With Other Networks

The Melbourne network is compared with seven other networks in Table 3. Three of those networks came from a division of Chase Manhattan Bank in New York (CMB) and three came from a dataset collected from regional districts of the Office for Civil Defense (OCD). In each of these organizations, data were collected on three topic areas: production, innovation, and maintenance. "Production" communication is communication about job-related matters. "Innovation" refers to communication about changes in organizational procedures, and "Maintenance" refers to social maintenance, or informal communication. The seventh network was collected from a naval installation on Treasure Island.

We proceed now with some comments on the numbers in Table 3.

1. The Number of Non-participants

In five of the networks the number of non-participants was over 50%. Farace and Johnson (1974) do not explain these high numbers in their paper which reports the data. When analyzing networks in the past, it has been common practice to use the following procedure: the

Table 2

THE BASIC FINDINGS OF THE MELBOURNE STUDY

Non-participants					
Number	Role	Number of Links	Links per Node	% of Nodes	% of Non-participants
23	Isolate T1	0	0	8.8	32.4
38	Isolate T2	38	1	14.6	53.5
10	Tree Node	22	2.2	3.8	14.1
73	Total	60	---	27.2	100.0

Participants						
Number	Role	Number of Links	Links per Node	% of Nodes	% of Links	Participants
12	Liaison	87	7.25	4.6	11.0	6.3
13	Other	48	3.69	5.0	6.06	6.8
25	Total Linker	135	5.4	9.6	17.0	13.1
113	Non-bridge group member	374	3.31	43.3	47.2	59.5
52	Bridge group member	283	5.44	20.0	35.7	27.4
165	Total group member	657	3.98	63.2	82.9	86.8
190	Total participants	792	4.168	77.8	100.0	100.0

BREAKDOWN OF ALL NETWORKS BY ROLES

Table 3

PARTICIPANTS: ALL NETWORKS

Participants for each network are broken down into "A. Linkers" (liaisons and others), and "B. Group Members." Linkers are subdivided into liaisons and others. Group members are subdivided into non-bridges (no bridge links) and bridges. Under both "Linkers" and "Group Members" there is a column of totals. Thus, under "Linkers" there are three columns: "1. Liaisons," "2. Others," and "3. Combined Liaisons and Others." Under "Group Members" there are "4. Non-bridge Members," "5. Bridge Members," and "6. All Group Members." Column 7 combines linkers and group members. For each column there are three numbers: "N" refers to the number of nodes in that category; "%T" refers to the percentage of the total network that is in the category; while "%P" is the percentage of participants that is in the category. For example, in the Melbourne Network, there are 12 liaisons (column 1), which comprise 4.6% of the total network and 6.3% of the participants in the network.

The bottom two rows summarize the table. For example, the numbers in column 2, row K, shows that, in the eight networks studied, 1.23% of all nodes, or 50 nodes, were others. While this is 1.23% of all nodes, it was 2.8% of participants. Row L, the bottom row, shows averages across all networks. Thus, on the average, 38.7% of all nodes, or 77.7% of participants, were group members. (Information in rows A,B,C,E,F,G computed from Farace and Johnson, 1974, Table 1.)

NON-PARTICIPANTS: ALL NETWORKS

Non-participants for each network have been broken down into: "1. Isolate T1," "2. Isolate T2," "3. Isolated Dyads," and "4. Tree Nodes." Column 5 has data for all non-participants combined. The last column, "Network Total," has the total number of nodes in the network. This number is equal to the sum of column 5 in this table and column 7 in the participant half of this table. Within each column, "N" is the number of nodes in that category; "%T" is the percentage of the total network that is in the category; and "%I" is the percentage of non-participants that is in the category. For example, in row j, column 1, there are 23 Isolates T1. This is 8.8% of the total network N, which is 261. It is also 32.4% of the non-participant population of the network, which is 71.

As in Table 3A and 3B, the bottom rows summarize the table.

Table 3

		A. Linkers								
		1. Liaisons			2. Others			3. Total		
		N	%T	%P	N	%T	%P	N	%T	%P
A	CMB Production	154	15.8	32.7	12	1.2	2.5	166	17.0	35.3
B	CMB Innovation	111	11.4	34.5	4	0.4	1.2	115	11.8	35.7
C	CMB Maintenance	69	7.1	21.2	12	1.2	3.7	81	8.3	24.8
D	Average All Chase	111.3	11.4	29.5	9.33	.933	2.47	120.7	12.37	31.9
E	OCD Production	45	18.8	22.2	7	2.9	3.4	52	21.7	25.6
F	OCD Innovation	24	10.0	25.8	1	0.4	1.1	25	10.0	26.9
G	OCD Maintenance	4	1.6	6.8	1	0.4	1.4	5	2.0	7.2
H	Average all OCD	24.3	10.1	18.3	3	1.2	1.97	27.3	11.2	19.9
I	Treasure Island	9	5.3	9.5	0	0	0	9	5.3	9.5
J	Melbourne	12	4.6	6.3	13	5.0	6.8	25	9.6	13.1
K	Sum All Networks	428	10.52	24.2	50	1.23	2.8	478	11.75	27.0
L	Average All Networks	53.5	9.32	19.75	6.25	1.93	2.09	59.7	13.97	21.1

		B. Group Members									7. All Participants	
		4. Non-bridge			5. Bridge			6. Total				
		N	%T	%P	N	%T	%P	N	%T	%P	N	%T
A	CMB Production	229	23.4	48.7	75	77	15.9	304	31.1	64.7	470	48.3
B	CMB Innovation	184	18.9	57.1	24	2.4	7.4	208	21.3	64.6	323	33.1
C	CMB Maintenance	204	21.0	62.6	41	4.2	12.6	245	25.2	75.1	326	33.5
D	Average All Chase	205.7	21.1	56.1	46.7	27.8	11.97	252.3	25.9	68.1	373	38.3
E	OCD Production	127	53.1	62.6	24	10.0	11.8	151	63.1	74.4	203	84.8
F	OCD Innovation	50	20.9	53.8	18	7.5	19.3	68	28.4	73.1	93	38.8
G	OCD Maintenance	56	23.4	81.2	8	3.3	11.6	64	26.7	92.7	69	28.7
H	Average All OCD	77.7	58.3	65.9	16.7	6.93	14.2	94.3	39.4	80.1	121.7	50.7
I	Treasure Island	7	4.1	7.3	79	46.5	83.1	86	50.6	90.5	95	55.9
J	Melbourne	113	43.3	59.5	52	20.0	27.4	165	63.2	86.8	190	77.8
K	Sum All Networks	970	23.85	54.8	321	7.89	18.1	1291	31.74	73.0	1769	43.50
L	Average All Networks	121.2	26.0	54.1	40.1	17.53	23.6	161.4	38.7	77.7	221.1	50.2

[contd.]

Table 3
[contd.]

			C. Non-participants								
			1. Isolate T1			2. Isolate T2			3. Isolated Dyad		
			N	%T	%I	N	%T	%I	N	%T	%I
A		Production	274	28.1	54.4	150	15.4	29.8	44	4.5	8.7
B	CMB	Innovation	379	28.9	58.3	174	17.8	26.8	40	4.1	6.1
C		Maintenance	393	40.4	61.0	151	15.5	23.4	42	4.3	6.5
D	Average All Chase		348.7	35.8	57.9	158.3	16.2	26.7	42	4.3	7.1
E		Production	12	5.0	33.3	21	8.8	58.3	2	0.8	5.5
F	OCD	Innovation	89	37.2	61.0	36	15.1	24.6	8	3.3	5.5
G		Maintenance	93	38.9	54.7	45	18.8	26.5	16	6.6	9.4
H	Average All OCD		64.7	27.0	39.7	34	14.2	36.5	8.7	3.6	6.8
I	Treasure Island		59	34.7	78.6	15	8.8	20.0	0	0	0
J	Melbourne		23	8.8	32.4	38	14.6	53.5	0	0	0
K	Sum All Networks		1322	32.5	57.6	630	15.5	27.4	152	3.7	6.6
L	Average All Networks		165.2	29.0	50.47	78.7	13.2	32.9	19	2.95	5.2

			C. Non-participants [contd.]					Network Total
			4. Tree Node			5. All		
			N	%T	%I	N	%T	
A		Production	35	3.6	6.9	503	51.7	973
B	CMB	Innovation	57	5.8	8.8	650	66.6	973
C		Maintenance	58	6.0	9.0	644	66.2	973
D	Average All Chase		50	5.1	8.2	599	61.6	973
E		Production	1	0.4	2.7	36	15.0	239
F	OCD	Innovation	13	5.4	8.9	146	61.0	239
G		Maintenance	16	6.6	9.4	170	70.9	239
H	Average All OCD		10	4.1	7.0	117.3	49.0	239
I	Treasure Island		1	0.6	1.3	75	44.1	170
J	Melbourne		10	3.8	14.1	71	27.2	261
K	Sum All Networks		191	4.7	8.3	2295	56.5	4067
L	Average All Networks		23.9	4.02	7.64	286.8	50.3	508.4

first submission to the computer usually includes all links. With many networks, the program produces either a very large number of "others" or else one single group to which everyone belongs. Neither of these outcomes are considered desirable. In order to obtain results in which most participants are members of relatively small, "clean" groups, weak links and unreciprocated links are dropped from the analysis. In the course of dropping these links, which are thought to be unreliable, many nodes are stripped of most or all of their links. This is probably the reason there are so many non-participants in these networks. Until the data collection methods are carefully studied, it will not be possible to understand the implications of dropping weak and unreciprocated links.*

The Melbourne network had only 27% non-participants, which is the second-lowest percentage of all the networks compared here. Nonetheless, it seems surprising that this many individuals would be so isolated from the rest of the system. The data that are available give no clues that could suggest why so many nodes are non-participants.

2. The Number of Liaisons and Others

Only 6.3% of participants in the Melbourne network were liaisons. This can be compared to the much higher percentage of the Chase networks (average = 29.5%), the OCD networks (average = 18.3%), and, to a lesser extent, the Treasure Island network (average = 9.5%).

---

*Farace and Johnson (1974) report breakdowns by percentage of all roles in their networks. Because so many non-participants were found in their networks, it was felt that it would be useful to report breakdowns of linkers and group members as percentages of participants, rather than as percentages of the whole network. When this is done, the variation in percentages within roles across networks is much lower, as can be seen in Table 3. For example, group members comprise from 21.3% to 63.2% of all nodes, but only 64.7% to 92% of participants.

Throughout the discussion of this chapter, most percentages and comparisons are made in terms of the number of participants or the number of participant-participant links. To insure clarity, however, it will always be noted which context is being used, unless it is obvious.

This difference is probably due to the fact that the individuals in the Melbourne system work in geographically separate areas, while both the Chase and OCD networks were from organizations that were totally contained in single locations.

3. The Number of Group Members

The proportion of nodes belonging to groups was higher in the Melbourne network (63.2% of all nodes) than in any other. With the exception of the OCD production network and the Treasure Island network (50.6%), no other network even comes close. These numbers should be qualified, however, with an awareness of the numbers of non-participants in the other networks. When proportions are calculated for participants only, the differences between highs and lows are much smaller. Here, 86.8% of Melbourne participants were members of groups, with the Chase networks averaging 68.1% and the OCD networks averaging 80.1%.

4. The Number of Bridges and Non-bridges

In both the Chase and OCD networks, most group members did not have bridge links to members of other groups. In most cases only one group member in five had any bridge links. In the Melbourne networ, the ratio of non-bridge to bridge was about two to one. In other words, one out of every three group members had bridge links to other groups.

The Individuals in the Network and Their Links

In this section we describe the individuals in the network, in terms of their linkage patterns. We begin with Table 4, which shows how many links each type of participant has. From this table it can be seen that role types differ markedly along this dimension. Liaisons have the most, followed by group members with bridge links and then group members who do not have bridge links.

Table 5 shows a similar comparison of roles across several networks. The main difference between the Melbourne network and the

Table 4

THE MELBOURNE NETWORK: THE LINKS OF PARTICIPANTS

		Number of Links by Roles: Participants						Grand Total Links
		Group Members			Linkers			
		Non-bridge	Bridge	Total	Liaison	Other	Total	
A	Number	374	283	657	87	48	135	792
B	Percent	47.2	35.7	82.9	11.0	6.06	17.0	100.0
C	Per node	3.31	5.44	3.98	7.25	3.69	5.4	4.168
D	# of nodes	113	52	165	12	13	25	190
E	E % links	59.47	27.37	86.84	6.32	6.84	13.16	100.0
F	E # links	471	216	688	50	54	104	792
G	S.D.	1.294	2.951	2.201	3.886	2.197	3.594	2.482

Row A shows number of links broken down by role category.
Row B shows proportion of total links by category.
Row C shows mean number of links per node by category.
Row D shows the number of nodes in each category.
Row E has the expected proportion of total links in each category, assuming even distribution across all nodes in network.
Row F shows expected number of links by category.
Row G shows S.D. of ℓ_1's for nodes in each category.

Table 5

MEAN NUMBER OF LINKS BY ROLES ACROSS NETWORKS

		Tree Nodes	Liaisons	Others	All Linkers	Non-bridge Group Members	Bridge Group Members	All Group Members	Total All Nodes*	Total All Participants
A	Chase Production	2.34	5.33	3.69	5.21	3.86	5.36	4.23	2.26	4.58
B	Chase Innovation	2.32	4.25	3.54	4.22	3.28	4.87	3.46	1.52	3.73
C	Chase Maintenance	2.24	3.92	3.06	3.79	3.52	4.17	3.63	1.52	3.67
D	OCD Production	2.00	6.28	4.71	6.07	4.62	6.12	4.86	4.23	5.17
E	OCD Innovation	2.30	5.00	3.71	4.95	3.18	1.44	2.72	1.70	3.32
F	OCD Maintenance	2.37	2.00	4.25	2.45	3.35	4.00	3.43	1.41	3.36
G	Melbourne	2.35	7.25	3.69	5.40	3.31	5.44	3.98	3.36	4.168
H	Average by Networks	2.27	4.86	3.81	4.58	3.59	4.49	4.04	2.29	4.000
G	Average by Nodes	2.31	4.91	3.68	4.78	3.65	4.86	3.89	2.00	4.14

*Including non-participants

This table compares the seven networks on the basis of the number of links each kind of node had. Row H shows the average numbers by networks. That is, to get the 2.27 in Row H, Column 1, the numbers in Rows A through G were summed and divided by 7. This average does not give larger networks greater weight.

Row I contains averages by node. Here, larger networks were weighted more so that each node, regardless of network size, counted equally in the calculation. Notice that there are slight differences between Rows H and I.

others is in the number of links the liaisons have. This might be explained by the physical differences in the structure of the organization -- the Melbourne system is scattered about a large geographical area, where all the other systems are located in single buildings. Liaisons would seem to be more important in holding the system together in the scattered system.

A slightly more detailed analysis is shown in Table 6, where the roles of both members of each linked pair have been counted. Points to notice include the following: (a) 792, or 90% of the total of 878 links, were between participants; (b) 516, or 65% of these 792 links, were within-group links; (c) all but 43 of the remaining 184 links connected either groups to other groups or groups to liaisons or others; and (d) most of the connections between groups were direct, without the intervention of liaisons or others.

The Groups in the Network

In this section we look at groups as units in themselves, rather than at their members or the links of their members.

1. How many and How Large

In Table 7 the group structure of the Melbourne network is compared to that of the Chase and OCD networks. It is clear from the numbers there that the Melbourne network, with a mean group size of 5.89 and a mean group density of 0.65, does not differ appreciably from the other networks.

Although these figures seem to indicate the stability of group sizes across systems, the validity of the stability cannot go unquestioned. We described earlier in this chapter the procedure used by investigators to obtain "good" results: if the number of others is too high or the number of groups too low, weak and unreciprocated links will be dropped until a "desirable" structural portrait of the system is obtained. If groups are either too few or too large, the data and the computer program will be manipulated until a "satisfactory" result is obtained. Thus, it is not clear whether the

Table 6

CROSS-ROLE CONNECTIONS
THE MELBOURNE NETWORK: BREAKDOWN OF ALL LINKS BY ROLES

	Within Group	Between Group	Liaison to Group	Other to Group	Total Group	Liaison to Liaison	Liaison to Other	Total Liaison	Other to Other	Total Other	Total Links Between Participants	Total Links in Network
Number	516	72	57	12	657	16	17	87	20	48	792	878
% of Total	58.8	8.2	6.5	1.4	74.8	1.8	1.9	9.9	2.3	5.5	90.2	100*
% of Column K	65.1	9.1	7.2	1.5	82.9	2.0	2.1	11.0	2.5	6.1	100.0	--

*The percentages do not add up to 100 because links to non-participants are not included.

The top row shows the number of links in each category. The middle row shows what proportion of the total linkage each category accounts for, and the bottom shows what proportion of links between participants each category accounts for.

Table 7

CROSS-NETWORK COMPARISON OF GROUP STRUCTURE

Network	Number of Groups	Number of Individuals in Groups	Mean Size	Range	S.D.	Mean Group Density*	Mean $\overline{D}$
Chase Prod.	46	303	6.58	3/18	3.31	0.644	1.472
Chase Innov.	39	208	5.33	3/11	2.62	0.691	1.377
Chase Maint.	44	245	5.57	3/18	2.90	0.683	1.393
OCD Prod.	17	151	8.88	3/18	5.47	0.620	1.471
OCD Innov.	13	68	5.23	3/9	2.04	0.702	1.385
OCD Maint.	9	64	7.11	3/11	2.76	0.534	1.735
Melbourne	28	165	5.89	3/16	3.18	0.653	1.443
Navy Expertise	27	272	10.1	3/*	*	0.50	*
Navy Expertise	44	407	9.3	3/*	*	0.57	*
Navy Social	16	162	10.1	3/*	*	0.60	*
Navy Social	38	415	10.9	3/*	*	0.53	*
Navy Authority	12	85	7.1	3/*	*	0.51	*
Navy Authority	30	275	9.2	3/*	*	0.44	*

*Data not available.

stability in group sizes is a real finding or an artifact of the analytical procedures used to understand the data.

2. Links in Groups

Most of the links of group members are with other group members. The remaining links are divided almost evenly between links with other groups and links with liaisons and others. The typical group has about four links tying it to the rest of the network -- two with other groups and two with liaisons.

3. Group Size, Group Density, and $\overline{D}$

While information about how the groups are related to other groups is important, it tells us nothing about linkage patterns within groups. To get this kind of information we need to shift our focus from the members of groups and their links to groups themselves. This has been done for the numbers plotted in Figures 1, 2, and 3. These plots show the relationships between group density, group size, and the mean distance from any member in the group to any other member.

We begin with Figure 1, which plots group size against group density. The relationship is clear: as group size increases, density decreases. This is not an unexpected finding -- the larger the group is, the more people there are for each person to be linked with. Since people are limited in the number of relationships they can sustain, they necessarily limit the number of links they have. Thus, it would be expected that in large groups the density would be lower than in small groups.

The numbers plotted in Figures 2 and 3 are based partly on the distance matrix for each group.* Figure 2 shows a plot of group

---

*In the distance matrix the entry in row i, column j is the number of steps in the shortest path from node I to node J. Thus, a direct link would be represented as a "1" in the distance matrix. A connection through one intermediary would be a "2" and so on. If the numbers in row i are averaged, the result is the mean number of steps it takes to reach any person in the group from node I. If all the row averages are themselves averaged, the result will be the group mean distance, $\overline{D}$, which is the average distance from any member of the group to any other member. [Footnote continued on page 229.]

Figure 1

THE RELATION BETWEEN GROUP SIZE AND GROUP DENSITY
IN THE MELBOURNE NETWORK

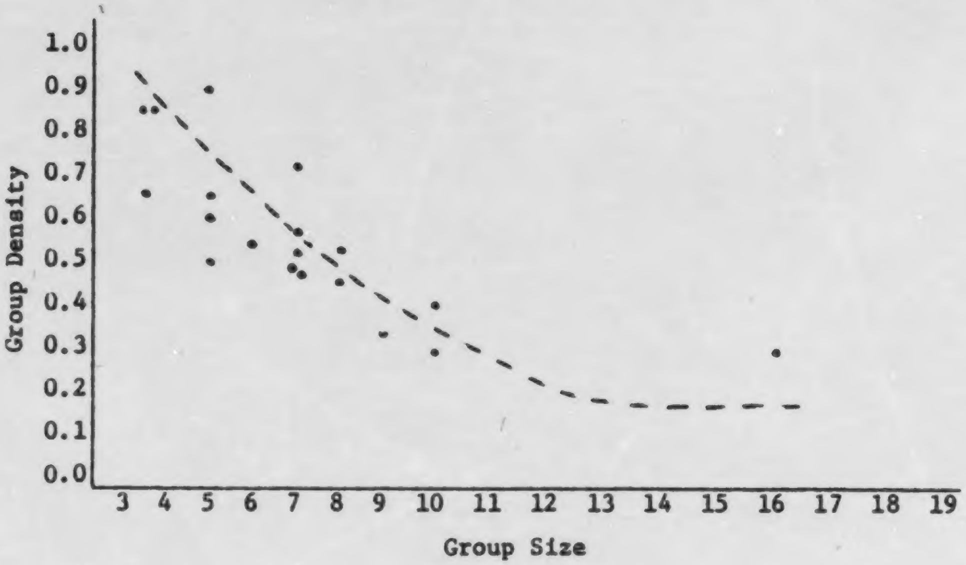


Figure 2

THE RELATION BETWEEN GROUP SIZE AND $\overline{D}$ FOR ALL GROUPS IN THE MELBOURNE STUDY

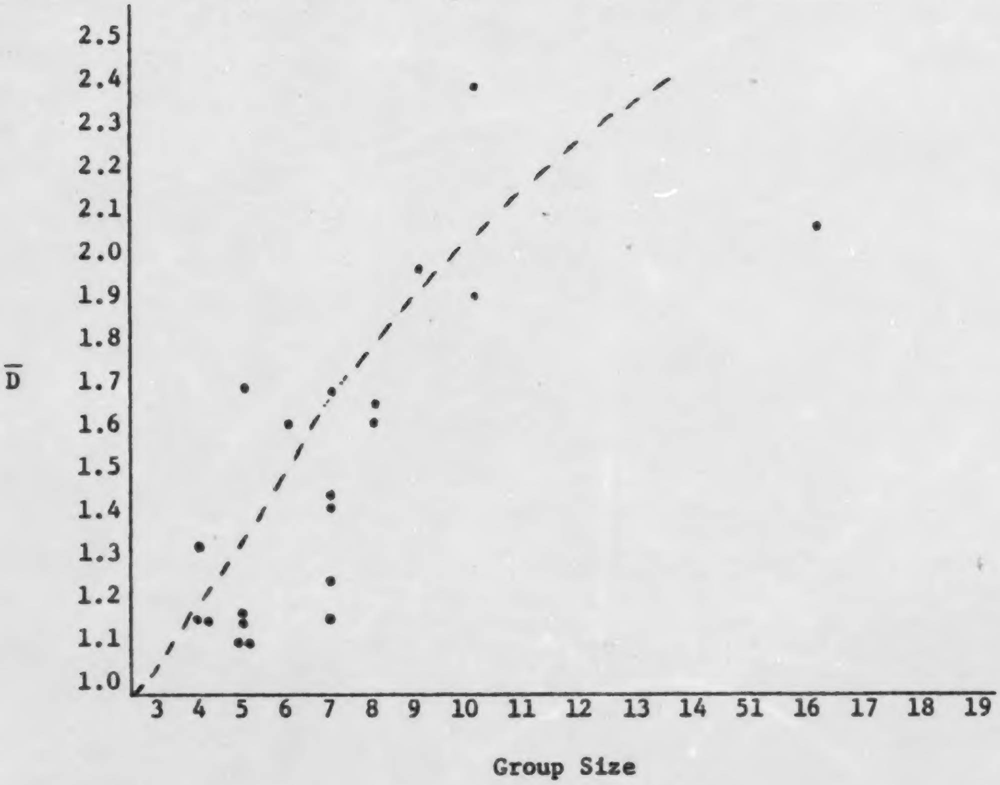
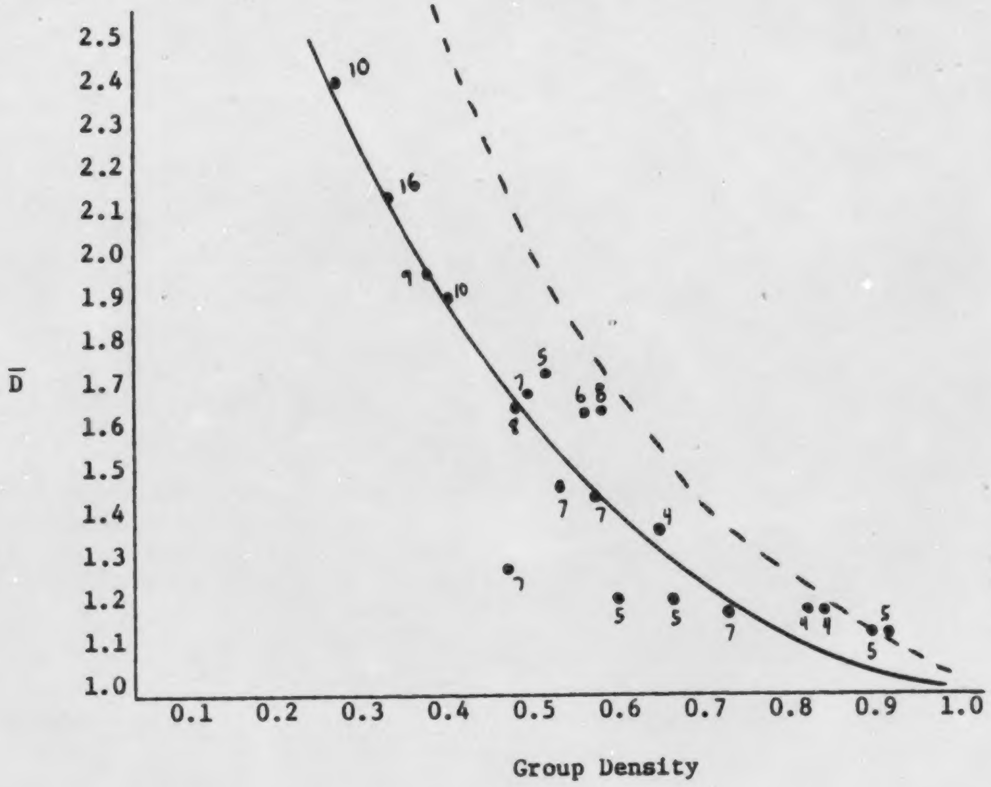


Figure 3

THE RELATION BETWEEN GROUP SIZE AND $\overline{D}$ FOR ALL GROUPS IN THE MELBOURNE STUDY (THE SMALL NUMBERS NEXT TO THE POINTS ARE GROUP N's)



size against $\overline{D}$, the average distance from any member of the group to any other member. The relationship observed here is roughly the opposite of the one shown in Figure 1, where group density was plotted against group size. Again, the relationship is not unexpected. The rationale here is the same one that led us to expect an inverse relation between group size and group density.

In Figure 3, $\overline{D}$ is plotted against group density. The small numbers next to the points refer to group size. A close examination of the plot will show that groups are roughly ordered by size along both axes of the plot, showing how group size is related to both density and $\overline{D}$.

4. Outside-group Connections

A relationship that is of interest is the nature of the nodes which connect the group to the rest of the system. We have already compared bridge group members with non-bridges, in terms of number of links. We might also ask if nodes that have connections outside the group are closer* (in terms of path length) to the rest of the group

---

[Footnote continued from page 225.]

If links are non-directed (symmetrical), the distance matrix will be symmetrical, and row means will equal column means. However, if links are directed (asymmetrical), rows will not necessarily equal columns. In this case, the mean of the entries in row i will be the average number of steps it takes node I to reach all the other group members. The column mean for node I will be the average number of steps it takes for all the other members to reach node I.

*The distance a node is from the rest of the group will vary as group size varies. To get an index of distance to the rest of the group that is free of contamination by group size, it was necessary to construct an index that was defined in terms relative to the other members of the same group. The index that was used in Figure 4 was calculated by dividing the average distnace from the node to other members by the overall group mean, $\overline{D}$. This gives a ratio that has a value of 1.0 when the node is the same distance as the group average, a value greater than 1.0 if the node is farther away, and a value smaller than 1.0 if the node is closer. For example, a distance ratio of 0.5 means that the distance from this node to the other group members is only half the average distance for the group. This value is insensitive to group size, and takes into account variations in group density as well.

than nodes who do not have outside connections. Figure 4 shows a comparison of nodes having one, two, three, four, and five connections outside the group.

As Figure 4 shows, the nodes with one outside-group link have distance ratios that scatter over a range from 0.6 to 1.4. The mean of these ratios is 0.99 -- a value less than 1.0, but not very convincing, given the variance of the distribution. This is roughly the same thing that is observed for nodes having two outside links. Nodes having three, four, five, or more outside links, however, have distance ratios that are considerably smaller. The more outside-group links there are, in fact, the lower is the distance ratio and the more consistent is the relationship. It would, then, seem that groups are structured in such a way that nodes having wide access to the rest of the network are closer to other group members than to nodes without such access.

This finding must be qualified, however, because it is possible that the relation is spurious. It might be that the nodes with many outside links are also the ones with many within-group links. If data were available from a greater number of groups, it would be practical to attempt to partial out the effects of number of links in some way that would allow the question of closeness to be answered with less equivocation. This task will be deferred to another time.

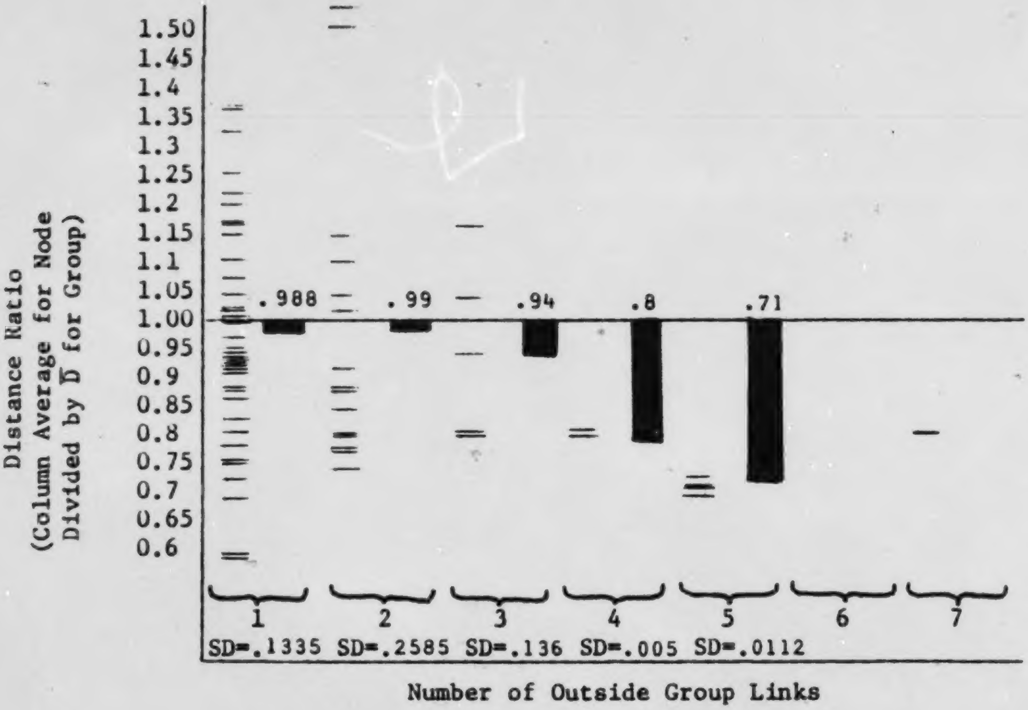
The System

In this section we shift our level of analysis to the whole system. Here we look at the system as composed of groups of individuals, together with the persons and links that connect those groups in the organization. For this analysis a pair of groups was considered to be linked if either (a) there was a bridge link between numbers of the two groups, or (b) both groups had links with the same liaison or other.

The results of the structural analysis are shown in Table 8. Because the variance in the number of links to each group was significantly greater than the expected variance, structural calculations were done with the equations that control for the distribution of links.

Figure 4

THE RELATION BETWEEN DISTANCE TO OTHER GROUP MEMBERS AND NUMBER OF OUTSIDE-GROUP LINKS



In this plot, the values on the ordinate were calculated by dividing the mean distance from the node to all other nodes by the overall distance for the group. If the node is closer to the other group members than all other group members, the result of the ratio will be less than 1.0. If the node is as far as the other nodes in the group, the ratio will be 1.0. If the node is farther than the other group members, the ratio will be greater than 1.0.

Each small line represents the distance ratio for one node.

The thick bars represent the mean distance ratio for all the nodes having each number of outside-group links.

For example, of the nodes having three outside-group links, the ratio of the distances from those nodes to other group members and the overall group mean distance averaged 0.94.

Table 8

STRUCTURAL ANALYSIS OF THE MELBOURNE NETWORK
USING GROUPS INSTEAD OF NODES

	EQN#**	Value	Description
L		69	Number of links
N		28	Number of groups
C	1	.18254	System density
C^2		.03332	
C^3	4	.00608235	
$\bar{\ell}$		4.9286	Mean of ℓ_i's
$S^2{}_o$	6	9.9949*	Observed variance in ℓ_i
$S^2{}_e$	5	4.029*	Expected variance in ℓ_i, given N and L
T_o		53	Observed number of triangles
T_e	2	19.92536	Expected T, given N and L
$T_{e.\ell}$	7	25.008	Expected T, given ℓ_i's
T_m	8	205.258	Maximum T, given N and L
T_{max}	3	3276	Maximum T, given N
$T_{m.\ell}$	10	137	Maximum T, given ℓ_i's
$S_{.\ell}$	12	.25	Structure, given N and ℓ_i's
S	11	.178	Structure, given N and L

*Observed differs from expected, significant at $p<0.01$ (F-test, df=260 and 260).

**All equation numbers are from Chapter Seven.

The results indicate higher structuring than would be expected by a random model; the conclusion is that the groups are organized in some non-random manner.*

A "sociogram" was drawn for the system, using the criteria for linkage outlined above. It was not possible to discern any differentiation into groups of groups from this graph, which is shown in Figure 5.

THE PURSUIT OF MEANING: RELATING NETWORK VARIABLES TO OTHER VARIABLES

While a familiarity with network characteristics is a necessary starting point, it is not, in itself, very satisfying. A more complete kind of understanding is gained by relating structural characteristics with other kinds of phenomena. In fact, it almost seems as if the only meaning that can be attached to network characteristics is meaning drawn from this additional kind of information. In this section we discuss three different approaches toward this greater understanding. The first compares networks for different types of communication and at different points in time for a single organization. The second examines some hypothesized antecedent and consequent variables of participation or isolation in networks. The third assumes group structure to be the antecedent variable and examines its impact on the attitudes of group members.

1. Structural Stability

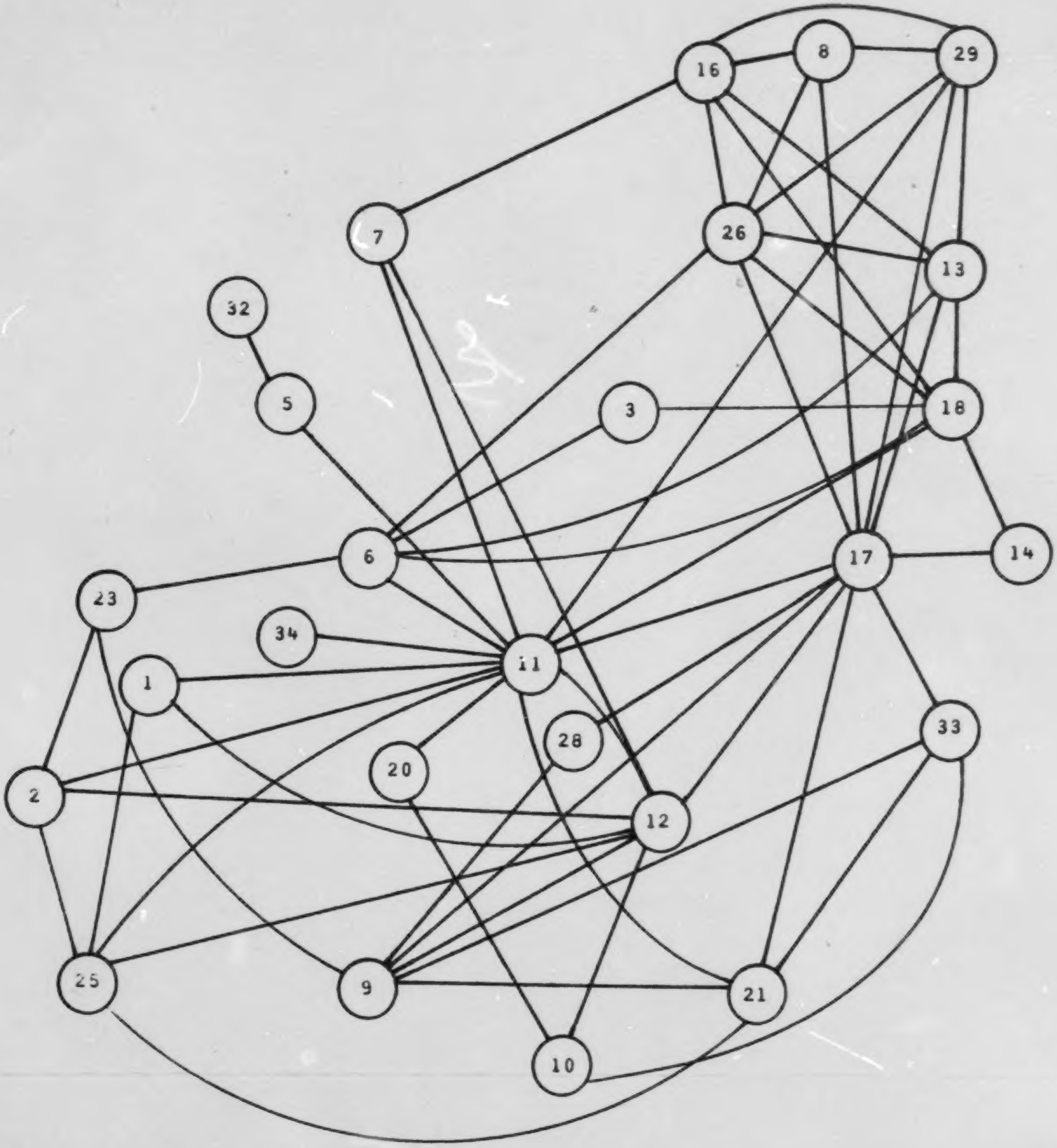
This study was done by Roberts and O'Reilly (1975a). (All references in this section are to those investigators.) They claim that there has been little research in the area of job-relevant

---

*There is a potential problem with this analysis in the way liaison and other links between groups were counted. Whenever two groups would each have links to the same liaison or other, they were considered to be linked. This would inflate the number of triangles by putting a link there even if no information were exchanged through the channel provided by the liaison or other. Although the procedure used is not ideal, it is the only one available at the time.

Figure 5

A "SOCIOGRAM" FOR THE GROUPS IN THE MELBOURNE NETWORK, DRAWN FROM THE MATRIX SHOWN IN FIGURE 10


16
8
29
7
26
13
32
5
3
18
6
17
14
23
34
11
1
33
28
20
2
12
25
9
21
10

correlates of differential individual communication role occupancy in organizations. Groups are most frequently described in terms of their size and the degree to which they are internally connected. There are no descriptions in the organizational literature of differential size and density of groups which develop for different content reasons. There is no indication, for example, that groups for a social network will be larger or smaller than groups for an expertise or an authority network (pp. 7-8). There is also little in the way of research on the stability of network structure over time.

The rationale, method, and subjects used in this study were described in an earlier section of this chapter. The relevant results are shown in Tables 9 through 13, and discussed in the following paragraphs.

Role occupation across time and content. Table 9 shows a breakdown of the members of this system into participants and isolates for the six networks. These numbers can be compared to the corresponding numbers for the Chase networks, the OCD networks, and the Melbourne network, which are also shown in Table 9. It is clear from this table that there is more variability in the Naval networks than in the others. Where the percentage of participants that were members of groups in the other networks ranged from 64% to 90%, in the Naval networks the range was from 25% to 100%. It is likely that this difference is due to differences in data collection procedures used by the investigators. Roberts and O'Reilly report a response rate of only 81%; the response rate in the other studies was always close to 95%. Another difference is in the way the sociometric questions were worded. While the other studies all used slight variations on the "Who do you talk with about ________" question, Roberts and O'Reilly had only one question that was comparable (the social network question). Their other two questions differed in two ways. The expertise question* asks for sources of information, and is clearly directional. The

---

*"When you need technical advice in doing your job, who are the persons you are most likely to ask?"

Table 9

COMPARISON OF PARTICIPANT AND ISOLATE PROPORTIONS OF 14 NETWORKS

Network		Participant	Non-participant	Group Membership	
Navy	Expertise 1	80.3%	19.7%	34.0%	42.3%
	Social 1	80.4%	19.6%	20.2%	25.1%
	Authority 1	42.1%	57.9%	10.6%	25.3%
	Expertise 2	73.4%	26.6%	56.1%	76.5%
	Social 2	72.2%	27.8%	57.2%	79.3%
	Authority 2	38.0%	62.0%	38.0%	100.0%
Chase	Production	48.3%	51.7%	31.1%	64.7%
	Innovation	33.1%	66.6%	21.3%	64.6%
	Maintenance	33.5%	66.2%	25.2%	75.1%
OCD	Production	84.8%	15.0%	63.1%	74.4%
	Innovation	38.8%	61.0%	38.4%	73.1%
	Maintenance	28.7%	70.9%	26.7%	92.7%
Melbourne		77.8%	27.2%	63.2%	86.8%
Treasure Island		55.9%	44.1%	50.6%	90.5%
		%T	%T	%T	%P

%T indicates that the proportions are a percentage of the total number of nodes in the network.

%P indicates that the proportions are a percentage of the number of participant nodes in the network.

authority question* is ambiguous as to direction -- it could mean either to whom do you complain (upwards) or to whom do you speak to rectify the situation (downwards).

A further difference that is related to the kinds of questions is the type of group structure that would be expected. Whereas the questions used by the investigators of the other studies all asked about communication concerning the sort of information that could flow horizontally through the network, both the expertise and authority questions in the Naval studies asked about information that would be expected to move vertically in a hierarchical network. One would expect groups in a horizontal network, but not in a vertical one. These problems all make comparisons between the different systems tenuous at best and invalid at worst.

In spite of the large differences noted above, the groups found in the Naval networks were surprisingly similar to those of the other networks. The figures were already shown in Table 7. The lower densities in the Naval groups are possibly due to the lower response rate of these studies. The stability of group sizes and densities across networks and across content areas is striking. Given the number of systems and their variety, it seems safe to make the general statement that groups tend to have between six and ten members, regardless of the topic or type of network.

Tables 10 and 11 allow a more powerful kind of statement to be made about structural stability. In Table 10, nodes were compared at two different times on the basis of their role in the network. In both the expertise and social networks over 70% of all nodes retained the same role at Time 2 that they had at Time 1. Table 11 shows similar comparisons between the different networks at one time.

Given the stability figures for the networks that had the measurement problems discussed above, it would be interesting to see what the same figures would look like when the problems were avoided. It would seem likely that stability would be higher in a more reliable

---

*"If you are upset about something related to the Navy or to your job, to whom in the squadron are you most likely to express your dissatisfaction (gripe) formally?"

Table 10

COMPARISON OF ROLES ACROSS TIME WITHIN NETWORK: THE NAVAL NETWORKS

Network	Role*	Time 1	Time 2	Stability of Role
Expertise	P	80.3%	73.4%	76.0%
	I	19.7%	26.6%	
Social	P	80.4%	72.2%	71.0%
	I	19.6%	27.8%	
Authority	P	42.1%	38.0%	59.0%
	I	57.9%	62.0%	

* P = Participant
I = Non-participant

Table 11

STABILITY ACROSS CONTENT AND TIME: THE NAVAL NETWORKS

	E	S	A
E	76.0%	81.9%	61.5%
S	80.4%	71.0%	60.0%
A	59.6%	58.7%	59.0%

Time 1

E = Expertise
S = Social
A = Authority

- The percentages below the diagonal are time 1 comparisons between different content networks.
- Time 2 comparisons are above the diagonal.
- The percentages on the diagonal are cross-time stabilities within content.
- Note the relatively low entries for all comparisons involving authority networks.

situation. If this were the case, the findings would have significant implications for managers of organizations who are concerned with the flow of information through the system.

2. Antecedents and Consequences of Network Role

This study was also conducted by Roberts and O'Reilly, using a combination of the network data discussed above and additional demographic and attitudinal data as well (this and all further references in this section are to Roberts and O'Reilly, 1975b). In a follow-up of their analysis of structure, the investigators examined "sets of demographic and intrinsic characteristics people bring with them to their jobs" to see if these "are related to communication role occupancy" (p. 1). Communication role occupancy was operationalized as either isolation or participation in the communication network. The role variable was also examined in relation to a set of individual responses thought to be "important in terms of organizational viability."

Besides the network instrument, respondents were asked to complete a survey instrument containing demographic questions and a series of questions concerned with their perceptions and feelings about various aspects of their work environments. The demograhic variables assessed were rank and tenure in the organization, population size of community in which the respondent was raised, amount of education, age, and tenure in the Navy. The instruments used to measure the perceptions and feelings about work environments are described in Roberts and O'Reilly (1975b, pp. 11-12).

Discriminant function analyses were used to differentiate isolates and participants across the three networks for the two sets of antecedent variables (demographics and intrinsic characteristics) and two sets of response variables (satisfaction and perceived communication).

(a) Results. The results of the discriminant functions were significant for both the social and expertise networks. In the expertise network, participants perceived that various communication modalities were used and that they (the participants) received redundant information. Isolates had an increased tendency to deliberately withhold information and report lower satisfaction with communication in general.

In the social network participants again felt they received redundant information. They also had a tendency to summarize to insure transmission of important information. Isolates perceived more use of written and telephone than did participants.

A general conclusion was that participation in communication is associated with perceptions of increased information flow, more redundancy, and greater overall satisfaction with communication. There were no significant differences in commitment to the organization between isolates and participants in either the social or the authority networks. However, participants in the expertise network were significantly more committed to the Navy than were isolates.

Differences in role occupancy and performance were observed with all networks. Participation in communication networks is associated with higher performance than is isolation. This difference is greatest for the expertise networks.

The general hypothesis was confirmed. Differential role occupancy is reflected in both antecedent and consequent conditions. Participation is generally associated with positive outcomes. Overall the picture is one of dysfunctional aspects for individuals who are not integrated into organizational communication networks.

3. The Impact of Group Structure on Attitudinal Configurations

Taylor (1976) took a slightly different approach in his study of the effects of group membership on the attitudes of the people in a network. He derived a theoretical rationale for the assumption that shared communication leads to a convergence of attitude from a general theory of attitude change provided by Woelfel and Saltiel (1975). This theory states that the attitudes a person has are a function of all the messages the person has received about the particular topics of interest:

> [The] consequent attitude *a* is the linear sum of the messages divided by the number *n* of messages .... Each message, x_i, is postulated to be a force which pulls the attitude in one direction or another .... Attitude change, then, is actually treated as a simple quantitative function of the number of messages an individual has received about a given attitude-object. (Taylor, p. 19, 1976)

Taylor hypothesized that communication groups would be more homogeneous with respect to a set of work-related attitudes than would non-group organizational members; that groups would differ from one another with respect to a set of work-related attitudes; that greater within-group connectedness would be associated with lower within-group variability on attitudes; that greater integrativeness would be associated with lower within-group variability; that greater connectedness would be associated with greater cohesiveness (1976, pp. 23-24).

The data employed in this study were collected from a Midwest manufacturing firm of moderate size (about 450 employees). The research was originally conducted in 1970 as part of an effort to monitor an ongoing organizational development program, which included the introduction of a Scanlon Plan some eighteen years earlier. Taylor comments that:

> It should be stressed that this organization is unusually sensitive to the need for cooperative or participative management systems. The resultant communication policy may be judged to have enhanced the amount of communication. (1976, p. 32)

Respondents were asked to fill out the Inter-Company Longitudinal Survey of the Institute of Social Research. Attitude items were measured on five-point Likert scales. Twenty-three items were chosen from the original set of 212 on the basis of a cluster analysis which showed that the 23 had a higher ability to discriminate between subjects. The original analysis of the data is described in Heinen (1972). In addition to the attitudinal items, respondents were asked to indicate the people with whom they interacted on the job. Content area was not restricted.

(a) Results. Because a one-way analysis of variance produced results that were not conclusive, a step-wise discriminant analysis was performed on the data with hierarchical clustering. This analysis identifies clusters of individuals on the basis of attitudinal similarities. The hypothesis was that the groups produced by the discriminant analysis would be the same as those produced by the network analysis -- demonstrating that the attitudes of the members of a group are similar to the attitudes of other members of the same group and different from the attitudes of members of other groups. Sixty percent of individuals were correctly placed by the discriminant analysis. Only 6% would be expected by chance. The finding was significant with $p<.0001$. The other hypotheses were either not supported or very weakly supported.

Taylor concluded that:

> The results provide strong support for the contention that communication groups are norm-providing mechanisms. The research demonstrates that unique constellations of attitudes can reliably recreate the informal groups detected by the network algorithm. Further, integrativeness and connectedness were found to be related to the process of the development of group normative structure .... The results show that members of informal groups feel less inclined to trust their work team members; less willing to advocate staying with the work team; less willing to rate highly the performance of their group; and, most importantly, less likely to see themselves as substantially important to the productivity of the team. (Pp. 55-56)

SUMMARY

Several uses of network analysis were reviewed. One network was examined in detail and the results interpreted in the context provided by the results of several other network analyses. To provide more meaning for the descriptive kinds of information provided by network analysis, three studies relating network characteristics to other phenomena were reviewed. The first related characteristics of networks for different content areas and across two points in time to obtain data on the stability of communication networks. The second examined some antecedents and consequences of communication network variables, and the third related attitudinal information to network structure.

CHAPTER ELEVEN

FURTHER

INTRODUCTION

We started Chapter One with a look at the basic models that have been used to structure approaches to the study of social systems. The first model we would recognize as taking a "scientific" approach was the mechanism--a machine obeying the laws of physics. This model was chosen because it was successful in other sciences and because it promised to move the social sciences closer to the more concrete hard sciences.

The fact that the mechanistic model, together with the classical analytic method which flowed from the assumptions of the mechanistic model, was failing to fit the reality of social systems became more and more obvious as more and more research was done using this approach. At this time, in the early nineteenth century, the organismic model started to grow in popularity as a replacement for the mechanistic approach.

Although it opposed the mechanistic approach, the organismic method was not noticeably more successful. In this ambiguous context, both approaches were used, sometimes consciously, and sometimes not (i.e., their assumptions were used, but not stated). All this time, though, there was a deep fundamental difference in the way the two models viewed the world. It was through this conflict that the systems approach was born.

The fact that the main concept in the systems approach is organization makes it easy to see why the systems approach was first used in the social sciences by people interested in the organization of social systems, which were viewed as communication networks.

We reviewed a number of the methods used to study organization in social systems in Chapter Two, and found most of them wanting. The most important lack, we said, was a lack of clear conceptual foundations. The whole concept of system had not been clearly enough explicated to allow formal analytic techniques to be drawn up. Furthermore, the processes of observation and description in the case of complex systems had not been looked into in sufficient detail to indicate how the relation between observer and system being observed limited the kinds of information that could be obtained in those situations. In Part Two, we looked into these problems and came up with a new approach to the study of complex systems.

We developed the model for the general case of complex multi-leveled systems. We started with an analysis of the logical form of these systems. There we were interested in exploring the logical implications of there being many levels in the system. The existence of multiple levels turned out to be centrally important for a number of different reasons. It meant that there were logical restrictions of many different types imposed on the system. These included restrictions on the way one system could interact with other systems, on the way the parts of the system could be related to the whole system, and on the characteristics of all systems organized in hierarchical fashion. When we examined the processes of description and observation in complex systems, we found that there were more implications of there being multiple levels in both the system being observed and in the observing system.

The processes of observation and description turned out to be much more complex than is generally recognized. Especially complex is the relation between the system being described and the different levels of descriptive statements that can be made about it. Over and over again, the concept of levels had to be considered in the analysis.

The concept of levels that played such a central role could not be understood without recourse to the more basic concept of constraint. Not only was constraint at the heart of organization, but also was it at the logical core of the process of description. This is why we spent so much time and energy trying to clarify the role of constraint in Chapter Five.

Part Three was organized in parallel to Part Two. While the beginnings of a conceptual framework were outlined in Part Two, the corresponding operational methods were sketched in in the chapters of Part Three. Much of what was described in those chapters was necessarily crude, as it represented only a preliminary outline. Much of what was described there has not been used enough in carefully evaluated research to let us know if it works. This is to be expected with newly developed methods.

In Part Four we presented some data. It was not nearly enough data to let us draw any conclusions. That was not the intent of the exercise. Rather, it was carried out to illustrate the kinds of analyses appropriate in the early stages of the work of the research program outlined in the earlier chapters. The numbers reported in Chapter Ten were not wholly fictitious. They were drawn from data sets that were chosen to give some valid impressions about some of the characteristics of the network aspects of the systems we are interested in understanding. As it is, the numbers only give some initial impressions -- they are in no sense comprehensive or complete descriptions. There is a lot more to our research program than network analysis. This additional work is the work of the future.

The State of the Art/Field/Area

An assessment of the strengths and weaknesses of the network approach will be a useful guide to the direction of future work. Several aspects can be identified for this type of analysis: guiding theory, statistical and methodological approaches, and the less interesting but more concrete research tools, such as data banks and computer programs.

As anyone who has actively engaged in network research will testify, the most advanced area is the third one mentioned above. The statistical and methodological foundations, while not entirely absent, are not as well-developed as the analytic methods.

The least developed of the three areas is the one that one would expect to come first -- that of guiding theory. How did this backwards situation come to be?

In the early stages of work on the problem of network analysis, it was easy to identify a specific, concrete need of researchers in the area. There was no fast easy way of digesting the sociometric data collected in network studies. The situation was one of a fairly clearly stated problem where the lack of answers seemed to be inhibiting progress in the general area. Instead of working on a better understanding of the theoretical issues involved, we chose to work on the seemingly more immediate problem of analytic techniques.

It may be unfortunate that we took this course. It is possible that the highly complex and sophisticated NEGOPY program will turn out to be a white elephant -- inhibiting further advancements in the very field that it was designed to help along. It is difficult to predict the kinds of needs that will go along with yet-to-be-advanced theory, but we attempt to do just that in the rest of this chapter.

WHERE WE ARE/WHERE WE WANT TO GO

The remaining sections of this chapter are organized in what now seems to be the correct order. Although hindsight suggests that theory should come first in research endeavors like ours, it does not seem to always be possible to advance the theories without having the kind of understanding that forays into methodological and analytical areas provide. We hope we will not be forced to backtrack too far. We begin the discussion with some comments on the sorts of theory we ought to have to guide research in the area. From there we move to a focus on the intermediate issues -- statistical and methodological considerations that flow from the coupling of theoretical biases and practical limitations. Finally, we turn to the concrete needs of those conducting research in the area.

Theory

In the chapters of Part Two we outlined a paradigm that suggests how we should view communication networks in complex systems. The approach we described there started with an analysis of the logical form of the systems we are interested in. We combined the results of this analysis with an analysis of the processes of observation and description to arrive at some statements about research goals. These statements described some of the characteristics of networks that should be important in the study of human communication systems — structuring, differentiation, roles, groups, and so on.

All these characteristics are endogenous to the network. They are based on distinctions made between one part of the network and the rest of the network, or between what was observed and what was possible.

A theory working with only this kind of information would be able to make advances in two directions. First, it would be possible to make statements about regularities observed in a variety of situations. This kind of research was reported in Chapter Ten, where we described such global phenomena as the "typical liaison" or the "typical group member" or the "typical group." Confidence is added to this kind of statement by looking at many networks in many different kinds of settings. This was one reason for the comparisons made between the Melbourne network and the others.

A second kind of statement that can be made within this kind of theoretical context would be the kind that related some endogenous network variables to other endogenous network variables. Here, instead of saying "the typical group member has four links," we would say things like "the more outside-group links a group member has, the closer that member is likely to be to the other members of the group." It is possible to make these kinds of statements with no data other than that which describes where the links are in the system.

The theoretical perspective we developed may be adequate, if we are satisfied with statements like the ones described above. If, on the other hand, we are interested in seeing how networks are influenced by other factors or how other factors are influenced by networks, we need a broader perspective. If we are to progress in a programmatic fashion, we need a perspective that will suggest whole families of research questions similar to the ones asked by Roberts and O'Reilly (1975a, b).

Much of the research taht has been done in this direction seems to be somewhere between the level of individual differences and system effects. For example, Roberts and O'Reilly were primarily interested in the behaviors and attitudes of persons who were either isolates or participants in the network. Taylor (1976) was interested in looking at attitudes of persons who were members of groups. Research relating properties of networks (i.e., properties of groups or properties of sets of groups) to endogenous variables is still rare.

One reason for the scarcity of research relating higher-level phenomena to exogenous factors is almost certainly the lack of guiding theoretical perspectives. It seems that this lack is not likely to be dealt with until the regularities of higher-level aspects of networks are better understood. Thus, the simpler studies relating endogenous variables to other endogenous variables seem likely to be more common in the near future. As a body of knowledge in this area is built up, more and more of the more complex studies relating endogenous higher-level network characteristics to exogenous factors will be completed.

If we are asked to describe the kinds of theory that are likely to be seen, we can only make the most general statements, since we do not understand even the kinds of things the theory is likely to relate. (If we had the theories today, we could dispense with the uninteresting content-free research into the nature of networks that makes up much of the work that has been done.) A starting point for this type of research has been suggested and some studies have been done. We refer here to studies looking at changes over time. If we are interested in dynamic processes, it seems fairly obvious that we will include change across time as one of the main concepts.

In the static one-point-in-time situation, we can observe some kinds of patterns or organization in the data. The most we can do with these observations is correlate them with other observations. We can make no statements about how things got to be the way they are, or how they influence other things around them. This is generally accepted to be the way science works.

When we include changes over time, we can observe the sequences and orders of changes. These observations allow causal statements to be made concerning the relations between phenomena. Instead of only being able to describe networks, we can begin to explain why and how they are the way they are.

Again, there has been little research in this area. The reasons for the lack here, however, are different from the reasons in other areas. The major problem here is that time-series studies are considerably more difficult to carry out than one-shot studies. The observational interventions are likely to be reactive -- systems do not permit themselves to be studied repeatedly -- and time-series studies are more expensive, as they demand more research assistant efforts and analytic complexity. They also take longer to do.

Statistical and Methodological Considerations

Closely related to the theoretical issues discussed above are problems related to statistics and methods. Four problem areas are discussed here: sampling, measurement, descriptive statistics, and inferential statistics.

1. Sampling

A theory of sampling for networks is yet to be proposed. Granovetter (1976) approaches the problem of sampling for a simple measure of a bsic network property -- density. It is unfortunate that he chose this measure to illustrate the problem, because density, as he operationalizes it, is an individual-level variable, rather than a network variable. The very crudity of the index makes it both easy to work with from a sampling perspective and of little interest from a network perspective.

The critical point that Grannovetter bypasses is that, when studying networks as networks, the network is the unit of analysis. Thus, one would sample networks, rather than individuals. When groups are the unit of analysis, one would sample intact groups -- not individuals or networks.

The problems with this approach are of two types. First, the statistics required to work with more complex indices are inordinately complex. Second, the actual operations of sampling -- i.e., locating and eliciting responses from networks or groups — gave yet to be worked out satisfactorily.

There is a third problem associated with the intact system approach: size. Many interesting systems are simply too large to be studied in toto. It seems that there are at least three ways of circumventing this problem. One is to divide the whole system into subsystems and to sample the subsystems. This would be the approach taken when a city was divided into neighborhoods and a subset of the neighborhoods was sampled for intensive study. The results would be generalized to the rest of the neighborhoods in the city. This method works when the system as a whole can be divided into "clean" subsystems — where the division does not do too much violence to either the subsystems or to the system as a whole. In many systems there are easily recognized subsystems that can be sampled in this way. Chavers' work with school systems (1976) illustrates this approach.

A second approach to the sampling problem is to use "snowball" techniques, perhaps starting each chain with a randomly selected individual. The statistics needed to make estimates of confidence with this method are formidable.

Benninger (personal communication) suggests a third method. He advocates dividing the network into a set of exhaustive and mutually exclusive subsets of individuals. For example, he might divide San Francisco into professionals, laborers, students, politicians, housewives, and criminals (assuming that these are mutually exclusive and exhaustive). He would then sample individuals from each category and study their links with individuals in other categories. With this information, he would be able to make statements about the way the individuals in any one category relate to the individuals in any other category. This "sociological" approach does not suffer from many of the problems that plague other schemes, but it provides less information than the others.

Clearly, the area of sampling is wide open and ready for some pathbreaking work.

2. Measurement

Closely related to the issue of sampling is the problem of measurement. Granovetter (1976) argues that problems with the recall method of collecting network data (see Chapter Six) are a major reason for developing some kind of sampling technique. His arguments apply more in very large systems where the task of recalling all one's contacts is too large for respondents to handle.

Measurement is not only a problem in very large systems, however. Even in moderately large (N=200 to 600) organizations, where the boundaries of the system are relatively clear, it is often difficult to get respondents to answer the fairly complex kinds of questions that must be asked to obtain useful network data. The discriminations required are complex and the data are thus not terribly reliable.* Not only are the data not that reliable, but also are they not that informative. It is difficult to get the amount of information that might be desired when the recall method is used. Diary methods, which can provide much more information, are also more obtrusive and meet more resistance on the part of respondents.

Finally, no studies have been done to clarify the meaning of some of the scales that have been used to code relationships in network studies. For example, an "importance" scale was used to rate each contact an individual had in some network studies. Importance scales usually run from 1 to 5, where 1 is not important and 5 is someplace on the other end of the continuum (see Chapter Six). It is not clear what kind of scale importance is. It probably is at least ordinal, but is it interval or ratio? Furthermore, do all individuals have the same meaning for "crucial" or do some consistently overrate the importance (relative to the others), while others underrate it? If there are differences among individuals, who are the over- and underraters? Can methods that control for this be developed? How important is the problem?

It seems clear that the use of sensitive analytic tools should be tempered with an understanding of just how uncertain the information in the data is. This area obviously is in need of more work.

---

*Recall the difficulties with the Roberts and O'Reilly data described in Chapter Ten.

3. Descriptive Statistics

When working with communication networks from complex systems it is absolutely necessary to have some way of digesting the massive amounts of information and producing concise and powerful descriptions of meaningful patterns in the data. This process involves the use of descriptive statistical summary statements which collapse large amounts of data into simple indicators of various characteristics. Two types of work need to be done in this area. The statistics we have need to be examined closely in a variety of situations to see how they behave, and new indicators need to be developed so that more can be said about the data we have.

Edwards and Monge (1975) did an empirical test of the relationships among several of the metrics that have been used to describe networks. In their analysis they were looking for sources of common and unique variance in each of the metrics. In this way they were able to reject several of the metrics as being redundant or vague, in terms of the large amount of variance they had in common with other, more basic, metrics. They were also able to identify metrics which shared almost no variance with others. These "unique" metrics were therefore demonstrated to be "clean," and not contaminated by other factors. More of this kind of commonality testing would give a clearer picture of the relationships among the other metrics. By sharpening our conceptual and operational terms, this process will allow us to make better use of existing indicators, as well as to discover the empirical "meaning" of any new ones that may be introduced.

The development of new indicators was mentioned as the other aspect that needs more effort. There are a number of indicators that are useful at the individual level of analysis. At higher levels there are fewer. If the higher levels are to be better understood, we need more indicators of different characteristics. One problem with higher-level indicators, however, is related not to other low-level indicators but rather to uncertainties in the measurement process. It is not clear, for example, what a link between two groups would look like. Other problems are caused by some aspects of the relationships that are not important at low levels but are crucial at higher levels. The problem of transitivity is one of these problems.* If this issue were better understood, it would be possible to develop a number of group-level and even system-level indicators that are not available at this time.

Analytic Considerations

Before the theories and statistics that will be developed can be put to work, actual techniques for doing this need to be provided. We have a start with the NEGCPY program described in Chapter Eight. Immediate attention can now be directed in four directions.

1. Improvements in Existing Programs.

There are many problems and limitations of the current NEGOPY program that must be dealt with.

(a) The existing program only runs on CDC 6000-series machines. It is imperative that a version compatible with IBM equipment be made available.

(b) The existing program is monolithically organized. Future versions should be modular, so that more poweful routines can be used without sacrificing core to code, and so that analyses can be done one step at a time -- giving the investigator both greater flexibility and greater economy.

(c) Much information produced by the program is not organized in a form that makes it readily available. More complete descriptive summaries should be included at many points in the analysis. When the program is organized in modular form, there will be room to include the code needed for these added features.

(d) The power of the program needs to be expanded to include multiple network or cross-time capabilities. Multiple network capabilities would allow several systems to be examined at once, with comparisons along a number of dimensions. Cross-time capabilities would allow changes in one network over time to be studied and possibly related to other, non-network variables. This step will be essential for theory building, at least as regards the dynamics of networks. Inputs to these routines would be the outputs of several of the earlier routines, rather than several raw data sets.

---

*See Chapter Six.

2. Development of Other Programs

While network analysis is an important part of this research program, it is only a part. As the sophistication of the approach grows, other types of analysis will be needed. If the complexity of these techniques are anything like that of others, computer programs will be necessary for carrying out any important research. These programs might be based on the logical model for complex systems that was described in Part Two. Alternatively, they might be based on dynamic theories of evolution and adaptation, *a la* Ashby (1956). The development of these new programs will mark the beginning of another surge in the advancement of the social sciences.

3. Moving Towards an "NPSS" (Network Package for the Social Sciences)

A major step in the development of modern statistical analysis was the presentation of a widely available package of standardized, powerful routines that performs most of the tasks that used to be done either by hand or by non-standard one-shot computer programs. A similar package of routines* in the systems/network area would likely have a major impact on the amount and quality of research done in the area.

DISCUSSION

The problems and tasks we have been discussing so far in this chapter are all important issues that will have to be considered during the next few years. Although they constitute a large research program, they only make the first step in the move toward a mature science of social systems. In this section we speculate on the directions such a science might take. These wanderings go back to the kinds of things we were discussing in the first and third chapters and move on from there.

---

*Perhaps a separate package would be less desirable than a merging of the two areas.

In the introductory remarks at the beginning we commented on a radical shift in communication styles that is influencing most of the world:

> This shift is fed by the continued development and advancement of new communication technologies. Thus, [for example] easy access to relatively inexpensive telephone equipment increase the amount of communication between distant areas .... As educational levels rise and political barriers drop, more and more people gain the ability to interact in the context of the emerging world society. Where in the past, local and national societies were forced to be independent of one another by a lack of communication facilities, they are now tied together into what is fast becoming a single integrated network of interdependent units, where the boundaries are becoming more and more mere political or economic considerations, instead of natural geographical or racial barriers.

We commented on the implications of this increased interdependence, which will include a whole set of new emergent properties -- both properties that couldn't exist before the communication ties and properties that are speeded up so that they become significant factors, leading to new changes that further accelerate the whole process. Situations in which a previously local event now has effects that shake the whole world system are becoming more and more frequent and undeniably important. It is obvious that the addition of more communication links is changing the fundamental nature of the system.

If we are to survive the turmoil of the coming years, which promises to change the world even more rapidly and dramatically than recent decades, we will have to have a better understanding of both the nature of the system and the nature of the kinds of changes we are likely to see. Without this understanding, we can only proceed blindly, accepting whatever consequences there are in store for us.

If we decide to approach the future with models that worked in the past, we cannot hope to succeed. The models that worked well in explaining and predicting a loosely connected and relatively stable world will not work in a volatile, tightly connected one. When communication and interaction increase, all the processes of change and growth speed up proportionally.

We described in this dissertation an approach that will move us closer to a science of large-scale social systems. The analysis presented here is only a beginning, however; it cannot be expected to provide even most of the answers. A more complete analysis would include many areas only touched upon here:

-- Perhaps most importantly, a highly developed theory and method for studying change in complex systems. What kinds of change can be observed? What causes change? What are the effects of change? How do we speak of change in quantitative terms? What kinds of change are good? What kinds are undesirable? Can we direct change? Can we avoid it? What kinds of systems (structures) change in the right ways? Can we alter system structure? What happens when systems disintegrate?

-- Theories and methods for approaching complex hierarchical systems. How do higher levels come to be? How do they influence lower levels? How are they influenced by lower levels? Can they be controlled/directed? What happens when a new level is added to the system? How can we study these levels? What can we do about them?

-- A theory of social system evoluation. How do complex social systems evolve over time? Are there parallels to biological evolution? Can this evolution be guided? If it can, what is the best way to move? Can we predict the direction of future evolutionary changes from the basis of present conditions?

-- Related to social evolution, a theory of adaptation and survival. How do social systems adapt to their environment? Are some kinds of systems inherently more able to adapt to changed conditions? If we can't direct the conditions, can we direct the ability of the system to adapt? Does the system survive in a changed state? Is social system survival a desirable thing? What are the impli-

> cations of survival? Can we predict which systems will survive? If adaptability insures survival in the short run, what does it imply in the long run? Are there other short-term/long-term tradeoffs?

Clearly the development of these theories and methods is a long-term program. The possible implications are unprecedented. We are at a critical point at this time. The new approach to social systems is in its early stages. It is especially fragile, and could easily be diverted away from the important long-term issues and toward more immediate short-term ones. The payoffs for doing short-term research are much more immediate and tangible than those for long-range work. This seems to be a danger that must be guarded against most strongly. Perhaps a compromise can be reached, where the attractive immediate applications can be done in such a way that the benefits can be funnelled back into the research effort in order to advance the long-range theoretical work.

CONCLUSION

We have here a beginning ....

BIBLIOGRAPHY

Ashby, H.R.
1956 *An Introduction to Cybernetics*. New York: Wiley.

Barnett, G.
1973 "The Nature of Random Stuctures." Michigan State University, E. Lansing, Michigan. Unpublished paper.

Barnlund, D.C.
1970 "A Transactional Model of Communication." In Sereno, K.K. and Mortensen, D.C. (eds.), *Foundations of Communication Theory*. New York: Harper & Row. Pp. 83-102.

Becker.
1968. "What Rhetoric (Communication Theory) is Relevant for Contemporary Speech Communication?" Paper presented at the Minnesota Spring Symposium in Speech-Communication.

Berlo, D.K.
1960 *The Process of Communication*. New York: Holt, Rinehart & Winston.

Berlo, D.K., Farace, Monge, Betty, and Danowski.
1972 *An Analysis of the Communication Structure of the Office for Civil Defense*. Technical Report, E. Lansing, Michigan State University, Department of Communication.

Beum, C.O. and Brundage, E.G.
1950 "A Method for Analyzing the Sociomatrix." *Sociometry* 13: 141-145.

Block, R.E. and Husain, S.Z.
1950 "An Adaptation of Holzinger's B-coefficient for the Analysis of Sociometric Data." *Sociometry* 13 (June): 146-153.

Borgatta, E.G. and Stolz, W.
1963 "A Note on the Computer Program for Rearrangement of Matrices." *Sociometry* 26: 391-392.

Breiger, Ronald L.
1976 "A Blockmodel Study of a Biomedical Research Network." *American Sociological Review* (February).
1974 "Duality of Persons and Groups." *Social Forces* 53,2: 181.

Breiger, R.L., Boorman, S. and Arabie, P.
1974 *An Algorithm for Blocking Relational Data, with Applications to Social Network Analysis and Comparison with Multidimensional Scaling*. Technical Report, Stanford University, Institute for Mathematical Studies in the Social Sciences (August).

Brown, G.S.
1969 *Laws of Form*. Great Britain: Allen & Unwin, Ltd.

Buckley, W.
1967 *Sociology and Modern Systems Theory*. Englewood Cliffs, N.J.: Prentice-Hall.

Chavers, D.
1976 Social Structure and the Diffusion of Innovations: A Network Analysis. Unpublished doctoral dissertation, Stanford University.

Coleman, J.S.
1964 Introduction to Mathematical Sociology. London: Free Press.

Coleman, J.S. and MacRae, D.
1960 "Electronic Data Processing of Sociometric Data for Groups up to 1,000 in Size." American Sociological Review 25,5: 722-727.

Collins, B.E. and Raven, B.H.
1969 "Group Structure: Attraction, Coalitions, Communication and Power." In Lindzey, G. and Aronsons, E. (eds.), The Handbook of Social Psychology IV (2nd ed.). Reading, Mass.: Addison-Wesley. Pp. 102-214.

Conrath, D.W.
1973 "Communication Environment and its Relation to Organizational Structure." University of Waterloo. Bell Northern Research Labs. Working Paper #69. January.

Dance, F.E.
1967 Human Communication Theory. New York: Holt.

Danowski, J.
1974 "Alternative Information Theoretic Measures of Television Messages: An Empirical Test." Paper presented to the Association for Education in Journalism, San Diego.

Deutsch.
1968 "Toward a Cybernetic Model of Man and Society." In Buckley, W. (ed.), Modern Systems Research for the Behavioral Scientist. Chicago: Aldine. Pp. 387-400.

Edwards, J.A. and Monge, P.R.
1975 "Descriptive communication Structure Metrics: A Preliminary Logical and Empirical Analysis." Paper presented to International Communication Association, Chicago.

Farace, R.V. and Johnson, D.
1974 "A Comparison of Selected Network Characteristics Across Six Organizations." Paper presented to annual meetings of International Communication Association, New Orleans.

Festinger, L.
1949 "The Analysis of Sociograms Using Matrix Algebra." Human Relations 2: 153-158.

Flament, C.
1963 Applications of Graph Theory to Group Structure. Englewood Cliffs, N.J.: Prentice-Hall.

Forsyth, E. and Katz, L.
1946 "A Matrix Approach to the Analysis of Sociometric Data: Preliminary Report." Sociometry 9: 340-347.

Garner, W.R.
1962 Uncertainty and Structure as Psychological Concepts. New York: John Wiley & Sons, Inc.

Gerard, H.
1968 "Units and Concepts of Biology." In Buckley, W. (eds.), Modern Systems Research for the Behavioral Scientist. Chicago: Aldine.

Granovetter, M.
1976 "Network Sampling: Some First Steps." American Journal of Sociology (in press).

Guimaraes, L.
1970 Network Analysis: An Approach to the Study of Communication Systems. Technical Report No. 12, Project on the Diffusion of Innovations in Rural Societies. E. Lansing, Michigan State University, Department of Communication (mimeograph).

Guttman, L.
1968 "A General Nonmetric Technique for Finding the Smallest Coordinate Space for a Configuration of Points." Psychometrika 33 (December): 469-508.

Hall, A.D. and Fagan, R.E.
1968 "Definition of System." In Buckley, W. (ed.), Modern Systems Research for the Behavioral Scientist. Chicago: Aldine. Pp. 81-92.

Harary, F., Norman, and Cartwright.
1965 Structural Models: An Introduction to the Theory of Directed Graphs. New York: John Wiley.

Heinen, S.
1972 "The Development of Working Teams in a Complex Organization." Unpublished doctoral dissertation, Michigan State University.

Holzinger, K.J. and Harman, H.H.
1941 Factor Analysis. Chicago: University of Chicago Press.

Hubbell, C.H.
1965 "An Input-Output Approach to Clique Identification." Sociometry 28 (December): 377-399.

Ilardo, J.A.
1973 "Ambiguity Tolerance and Distorted Communication: Therapeutic Aspects." Journal of Communication 23,4: 371-391

Jacobson, E. and Seashore, S.E.
1951 "Communication Practices in Complex Organizations." Journal of Social Issues, 7,3: 28-40.

Jacob, M.A.
1971 "The Structure and Functions of Internal Communication in Three Religious Communities." Unpublished doctoral dissertation, Michigan State University, 1971.

Katz, L.
1947 "On the Matrix Analysis of Sociometric Data." Sociometry 10:233-241.

Kruskal, J.B.
1964 "Multidimensional Scaling: A Numerical Method." Psychometrika 29 (June): 115-129.

Lankford, P.
1974 "Comparative Analysis of Clique Identification Methods." Sociometry 37: 287-305.

Lasswell, H.D.
1948 "The Structure and Function of Communications in Society." In Bryson L. (ed.), The Communication of Ideas. New York: Harper & Row. Pp. 37-51.

Lingoes, J.C. and Roskam, E.
1971 A Mathematical and Empirical Study of Two MDS Algorithms. Michigan Mathematical Psychology Program 71-1.

Lorrain, F. and White, H.C.
1971 "Structural Equivalence of Individuals in Social Networks." Journal of Mathematical Sociology 1: 49-80.

Luce, D.R. and Perry, A.D.
1949 "A Method of Matrix Analysis of Group Structure." Psychometrika 14: 95-116.

MacDonald, D.
1970 "Communication Roles and Communication Content in a Bureaucratic Setting." Unpublished doctoral dissertation, Department of Communication, Michigan Stata University, E. Lansing, Michigan.

MacRae, D., Jr.
1960 "Direct Factor Analysis of Sociometric Data." Sociometry 23 (December): 360-371.

McCroskey, J.C.
1968 An Introduction to Rhetorical Communication. Englewood Cliffs, N.J.: Prentice-Hall.

McGee, V.
1968 "The Multidmensional Scaling of N-sets of Similarity Measures: A Nonmetric Individual Differences Appraoch." Multivariate Behavioral Research 2 (September): 233-248.

McQuitty, L.L.
1957 "Elementary Linkage Analysis for Isolating Orthogonal and Oblique Types and Typal Relevancies." Educational and Psychological Measurement 17 (September): 207-229.

Miller, G.A.
1968 "What is Information Measurement?" In Buckley, W. (ed.), Modern Systems Research for the Behavioral Scientist. Chicago: Aldine.

Monge, P.R., Kirste, K.K., and Edwards, J.A.
1974 "A Causal Model of the Formation of Communication Structure in Large Organizations." Paper presented to annual meeting of International Communication Association, New Orleans.

Monge.
1973 "Human Communications: The Systems Perspective." Unpublished doctoral dissertation, Department of Communication, Michigan State University, E. Lansing, Michigan.

Moreno, J.L.
1934 Who Shall Survive? Washington, D.C.: Nervous and Mental Disease Monograph No. 58.

Morton, A.S.
1959 Similarity as a Determinant of Friendship: A Multidimensional Study. Princeton: Educational Testing Service.

Nagel, E.
1961 *The Structure of Science*. New York [Burlingame]: Harcourt, Brace & World, Inc.

Pattee, H.H.
1973 *Hierarchy Theory*. New York: Braziller.

Parker, E.
1975 *Social Implications of Computer/Telecommunications Systems*. Report #16, Program in Information Technology and Telecommunications, Center for Interdisciplinary Research, Stanford University (February).

Porat, M.
1975 *Defining an Information Sector in the U.S. Economy*. Report #15, Program in Information Technology and Telecommunications, Center for Interdisciplinary Research, Stanford University (February).

Proctor, C.H. and Loomis, C.P.
1951 "Analysis of Sociometric Data." In Jahoda, M., *et*. *al*. (eds.), *Research Methods in Social Relations*, Volume II. New York: Dryden. Pp. 561-586.

Rapoport, A.
1968 "Foreword." In Buckley, W. (ed.), *Modern Systems Research for the Behavioral Scientist*. Chicago: Aldine. Pp. xiii-xxii.

Richards, W.D.
1971 "An Improved, Conceptually Based Method for the Analysis of Communication Networks in Large Complex Organizations." Paper presented to International Communication Association, Phoenix.
1974 "Network Analysis in Large Complex Systems: Metric." Paper presented to International Communication Association, New Orleans.

Roberts, K. and O'Reilly, C.
1975a "Organizations as Communication Structures: An Empirical-theoretical Approach." Technical Report on ONR Researc Project. University of California, Berkeley.
1975b "Communication Roles in Organizations: Some Potential Antecedents and Consequences." Technical Report on ONR Research Project. University of California, Berkeley.

Roby, T.B.
1968 Small Group Performance. Chicago: Rand-McNally.

Rogers, E.M., Park, H.J., Chung, K.K., Lee, S.B., Puppa, W.S., and Doe, B.A.
1975 "Network Analysis of the Diffusion of Family Planning Innovations Over Time in Korean Villages." Paper presented to Population Association of America, Seattle (April).

Russell, H.
1974 "Communication Network Study." Memorandum to Department of Agriculture, Victoria, Australia (October 11).

Russell, B.
1971 *Introduction to Mathematical Philosophy*. New York: Simon and Schuster.

Shannon, E.C. and Weaver, W.
1949 *The Mathematical Theory of Communication*. Urbana, Ill.: University of Illinois Press.

Shepard, R.C.
1962a "The Analysis of Proximities, I." *Psychometrika* 27 (June): 125-140.
1962b "The Analysis of Proximities, II." *Psychometrika* 27 (September): 210-246.

Simon.
1973 "The Organization of Complex Systems." In Pattee (ed.), *Hierarchy Theory*. New York; Biaziller.

Sladek, J.T.
1968 *Mechasm*. New York: Ace. [Original title: The Reprodutive System]

Taylor, J.
1976 "Normative Referential Function of Informal Groups in Organization." Unpublished master's thesis, Department of Communication, Michigan State University.

Torgerson, W.S.
1958 *Theory and Methods of Scaling*. New York: John Wiley.

Toulmin, S. and Goodfield, J.
1962 *The Architecture of Matter*. New York: Harper & Row.

Tyron, R.C. and Bailey, D.C.
1970 *Cluster Analysis*. New York: McGraw Hill.

Weiss, R.S.
1956 Processes of Organization. Ann Arbor, Mich.: Institute for Social Research, University of Michigan.

Weiss, R.E. and Jacobson, E.
1955 "A Mathod for the Analysis of the Structure of Complex Organizations." *American Sociological Review* 20: 661-668.

Westley, B.H. and MacLean, M.
1957 "A Conceptual Model for Communication Research." *Journalism Quarterly* 34: 31-38.

White, H.C.
1974a "Models for Interrelated Roles from Multiple Networks in Small Populations." In Knapp, P.J. and Meyer, G.H. (eds.), *Proceedings of the Conference on the Application of Undergraduate Mathematics in the Engineering, Life, Managerial, and Social Sciences*. Atlanta: Georgia Institute of Technology.
1974b "Multiple Networks in Small Populations: II. Compound Relations and Equations." Unpublished paper, Harvard University, Department of Sociology.
1970 *Chains of Opportunity*. Cambridge, Mass.: Harvard University Press.
1963 *An Anatomy of Kinship*. Englewood Cliffs, N.J.: Prentice-Hall.
1961 "Management Conflict and Sociometric Structure." *American Journal of Sociology* 67: 185-199.

White, H.C. and Breiger, R.L.
1973 "Multiple Networks in Small Populations: I. Blockmodels." Unpublished paper, Harvard University, Department of Sociology.

Woelfel, J.
1970 "Sociology and Science." Unpublished paper, University of Illinois, Department of Sociology.

Woelfel, J. and Saltiel, J.
1975 "Cognitive Processes as Motion in a Multidmensional Space. A General Linear Model." Paper presented to the annual meeting of the International Communication Association, Chicago.

APPENDIX: NEGOPY, the Network Analysis Program

Documentation

A manual describing the network analysis program is available from the Institute for Communication Research, Stanford University, Stanford, CA 94305. The document is entitled "A Manual for Network Analysis, Using the NEGOPY Program." The author is William D. Richards, Jr.

There are sections in the manual describing the goals of network analysis, the collection of data, instrumentation, data preparation, and data analysis. The algorithm upon which the NEGOPY program is based is described. The parameters governing the operation of the program are discussed, with examples showing the implications of different options. Finally, for each section, the output of the program is described so that it can be easily interpreted by the user.

Availability

The NEGOPY program was written in CDC FORTRAN EXTENDED. The code is highly machine dependent; it runs only on CDC 6000-series machines. The program itself occupies about 3,000 cards in source version. The object version occupies roughly 27K (octal) words of core (60-bit words). Further information about the organization or availibility of the program can be obtained from William D. Richards, Jr. Institute for Communication Research, Stanford University, Stanford, CA 94305.